CONTENTS *(continued)*

The Study of Variability in Engineering Design *S. James Morrison*

A Sample Survey with both a Postal and an Interview Stage *Percy G. Gray*

Notes and Comments .. ..

Meetings of Sections of the Royal Statistical Society .. .. ..

Book Reviews:

Statistics: a New Approach, W. Allen Wallis and Harry V. Roberts

Statistics and their Application to Commerce, A. R. Ilersic .. ..

Guide to Elementary Statistical Formulas, Robert E. Johnson and Doris N. Morris

Publications Received .. ..

No. 3. NOVEMBER

The Sampling of Road Traffic .. .. .. *John C. Tanner*

Normality in Quality Control Charts .. .. .. *Peter G. Moore*

Differences between Experienced Interviewers .. .. *Alexander Barr*

'Three-Group' Regression Analysis. Part II. Multiple Regression Analysis .. .. *Wendy M. Gibson and Geoffrey H. Jowett*

A Simple Sequential Procedure to test whether Average Conditions achieve a Certain Standard .. .. .. *Rolfe C. Tomlinson*

An Optical Periodograph for Use in the Analysis of Continuous Measurements .. .. *Kenneth W. Kemp*

Confidence Limits for the Range of a Search Radar .. *Colin Stevens*

Notes and Comments .. ..

Meetings of Sections of the Royal Statistical Society .. .. ..

Book Reviews:

National Income Visualized, Arthur O. Dahlberg .. .. ..

Statistical Methods and Scientific Inference, Sir Ronald A. Fisher ..

A Dictionary of Statistical Terms, Maurice G. Kendall and William R. Buckland

The Lognormal Distribution, J. Aitchison and J. A. C. Brown ..

Statistical Methods in Research and Production, O. L. Davies (Editor)

Sudan Population Census Reports .. ..

Publication Received .. ..

CONTENTS

No. 1. MARCH

		Page
e Scope and Limitations of Econometrics	*L. R. Klein*	1
e Effect of obtaining Information from Different Informants on a Family Morbidity Inquiry	*Ann Cartwright*	18
ear Programming in Practice	*Emil Kay and Eric Duckworth*	26
npling for Variables with a Very Skew Distribution	*R. Tulse*	40
alysis and Presentation of the Results of Factorial Experiments	*Anthony G. Baker*	45
Problem in Production Control	*Raymond Chittleburgh and Alec Lee*	56
e of Regression Analysis to detect Errors in Measurement of Intermediate Materials in a Multi-stage Process	*William Stannage*	63
lculating the Standard Error of a Proportion	*Edward F. Brayer*	67
tes and Comments		69
etings of Sections of the Royal Statistical Society		71
ok Reviews:		
Experimental Design and its Statistical Basis, D. J. Finney		73
Guide to New Zealand Official Statistics, E. P. Neale		74
Curso de Estadística, Vol. 1, Enrique Chacón		74
Copybook for Beginners in Research Work, Nellie Landblom		75
A Primer of Social Statistics, S. M. Dornbusch and C. F. Schmid		75
Dynamic Factors in Industrial Productivity, Seymour Melman ..		76
The Classification and Coding of Accounts, J. M. S. Risk		78
The Essentials of Educational Statistics, Francis G. Cornell		78
The Theory of Games and Linear Programming, S. Vajda		79

No. 2. JUNE

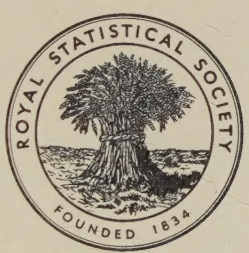
olutionary Operation: A Method for increasing Industrial Productivity	*George E. P. Box*	81
ne Statistical Applications of Electronic Computers	*Stephen Lipton*	102
ree-Group' Regression Analysis. Part I. Simple Regression Analysis	*Wendy M. Gibson and Geoffrey H. Jowett*	114
criminant Functions of Socio-Economic Class	*Kathleen Gales*	123

APPLIED STATISTICS

A Journal of the
Royal Statistical Society

EDITOR
DONALD G. BEECH

ASSISTANT EDITOR:
BERNARD T. RAMM



VOLUME SIX

Oliver and Boyd Ltd.
London: 39A Welbeck Street, W.1
Edinburgh: Tweeddale Court

1957

Digitized by the Internet Archive
in 2025

https://archive.org/details/journal-of-the-royal-statistical-society-series-c_1957_6

PLIED
ATISTICS

HE SCOPE AND LIMITATIONS OF ECONOMETRICS

L. R. KLEIN

Dr Klein first considers in general the way in which econometric techniques can be developed to estimate the interrelationships among economic variables, and then deals in some detail with applications of econometrics to particular problems.

Meaning of Econometrics

would be misleading to typify *econometrics* by the range of contribu- to *Econometrica*, the journal of the Econometric Society. The es in that journal cover a range wider than the theory and appli- n of measurement in economics, a plausible general definition of ometrics. Measurement in economics can, however, cover all of tical economics and indeed all numerical calculations in eco- cs. By econometrics we shall, in this paper, think of measurement d by some underlying mathematically expressed model. The urement aspect is, however, central. We want to rule out pure ematical economics, for that simply embraces economic theory in ticular form.

e theory of econometrics tries to discover how best to measure in omics, i.e. it sets out the properties of measurement methods so one is able to choose among alternative techniques. The applica- of econometrics attempt to carry out these measurements using l data gleaned from the economic process.

e statistical problems of econometrics are simply problems of ematical statistics as we generally understand that subject, but is one dominating characteristic of the econometric problems, ly, that they deal with *non-experimental* data. Haavelmo has put latter neatly in observing that 'Nature makes experiments for us'. inciple, it may be possible to conduct controlled experiments in omics and collect data from these laboratory situations but, except ossible rare cases in the future, the econometrician must accept ever data he can collect from uncontrolled observations of actual mic life. These considerations make the statistical tools of econo- cs somewhat distinctive in appearance.

ical Problem

lowing the great ideas of Walras, we view the economic system pable of being described by a system of simultaneous mathe- al equations. In actual practice we shall, of course, drop the nously refined detail of the Walrasian system but retain the main

idea of picturing the economy as a model written in the form system of mathematical equations. Our problem then will be to mate the parameters of these equations. The problem of estim the interrelationships among economic variables is not by any m the sole problem of econometrics, but it is, perhaps, the most con problem.

The mathematical systems to be estimated are not simply stun upon or grabbed out of the air; they are carefully thought out o basis of *a priori* information—information that is independent o particular sample being used. The *a priori* information may cons economic theory, a knowledge of economic institutions, a know of technology, or empirical results from independent samples. far as abstract economic theory is built into the mathematical m the econometric approach provides a systematic test of theory.

In order to describe the set-up of the econometric problem base the mathematical interrelationships of the economy, it is useful to tinguish two types of variable—endogenous and exogenous. Intuit these are economic and non-economic variables. They can, how be given a more rigorous definition in terms of probability distribut

Denote the ith endogenous variable by y_{it} and the ith exoge variable by z_{it}. The economic structure consists of a set of stoch equations relating the y_{it} to the z_{it}. The stochastic variables are ran disturbances, u_{it}, introduced in each equation. In addition the genous variables may be mutually interrelated by another stoch model. The equations of the system can be written as

$$\begin{aligned} f_i(y_{1t}, \ldots, y_{nt}, z_{1t}, \ldots, z_{mt}) &= u_{it} \qquad i = 1, 2, \ldots, n \\ g_j(z_{1t}, \ldots, z_{mt}) &= v_{jt} \qquad j = 1, 2, \ldots, m \\ & \qquad\qquad t = 1, 2, \ldots, T \end{aligned}$$

The f_i-functions are the structural equations which the econometr tries to estimate from sample data. The g_j-functions are not pa the econometric model as such. The disturbances u_{it} may be sidered as agglomerations of numerous minutiae which are too s to be treated separately. Together these additional variables ha noticeable effect on the economy, but we say no more about them can be inferred from the estimates of their assumed probability distr tion. In other words, chance takes care of our ignorance. It i important feature of modern econometrics that we do not approa problem with a strong predisposition that the u_{it} are, in some s small, i.e. we are not simply searching for high correlations. We trying to ascertain something about the structure of the f_i-func and the probability distribution of the u_{it}. In many cases we may be able to measure or otherwise collect explicit data on all the rele endogenous and exogenous variables. The omitted factors are s how included in u_{it}, and we assume that whatever these variable they follow some well-defined probability law. In linear system naturally assume that the u_{it} are linear combinations of indepen

bles which, by the central limit theorem, tend to be normally buted.

ıe exogenous variables, z_{it}, could, as in the usual regression model, set of known numbers not following the laws of probability. In ıodel above we have assumed that the z_{it} are interrelated by their set of equations. For example, the most perfect set of exogenous bles are climatic and weather factors. They have a one-way ıtive influence on the system. They are, on the other hand, rated by their own stochastic mechanism. The g_j-functions are, speak, the laws of meteorology.

l variables in the system are given the subscript t, possibly and ably to denote time values. We need not, however, deal exclu- with time series. The t-subscripts could equally well represent ally distinct points at some instant of time. The sample being used sts of T observations. We also note that among the set of exo- us variables we could include values of the endogenous variables ng to a prior time-period. These lagged variables are not exo- us but may be so regarded for statistical purposes, especially if T ge.

ıe econometrician starts out with the joint probability density of ver all sample values:

$$\mathrm{p}(u_{11}, \ldots, u_{nT})$$

equations of the economic system are regarded as transformation ions carrying the u_{it} into the y_{it}. The conditional density function e y_{it} given the z_{it} is obtained from the usual formula for trans- ing variables in probability distributions.

$$\mathrm{p}(y_{11}, \ldots, y_{nT} | z_{11}, \ldots, z_{mT}) = \mathrm{p}(u_{11}, \ldots, u_{nT}) \left| \frac{\partial(u_{11}, \ldots, u_{nT})}{\partial(y_{11}, \ldots, y_{nT})} \right|$$

$$= \mathrm{p}(\mathrm{f}_1, \ldots, \mathrm{f}_n) \left| \frac{\partial(u_{11}, \ldots, u_{nT})}{\partial(y_{11}, \ldots, y_{nT})} \right|$$

original density function must be multiplied by the Jacobian of ransformation in order to obtain the density function for the new bles, those of direct economic significance. The joint probability e endogenous variables is thus a function of the parameters of the ıctions. The problem reduces to the conventional problem of tical inference—to make probability statements about the para- rs of a universe on the basis of a sample. From the joint distribu- of the v_{jt} we can derive the joint distribution of the exogenous bles, and if this distribution is independent of the parameters of e are justified in ignoring the probability aspect of the exogenous bles or, in other words, can treat them as a set of known numbers. enous variables can thus be defined as those whose joint proba- distribution is independent of the parameters in the joint ability distribution of the endogenous variables in the model being dered.

Most readers are more familiar with the linear case and, indeed, theory is most fully developed for that case. If the f_i-functions v linear (or simply transformed into a linear set) and if the disturba were normal we should almost have before us the familiar proble maximising the likelihood of the sample by minimising a sum squares or, more generally, a quadratic form. The only depar from that straightforward problem would be the presence of expression

$$\left|\frac{\partial(u_{11}, \ldots, u_{nT})}{\partial(y_{11}, \ldots, y_{nT})}\right|$$

in the derived distribution of the endogenous variables. For the lin case this Jacobian expression would simply be the determinant of coefficient matrix of the system.

In some special cases this approach to the econometric problem be greatly simplified. If the Jacobian expression is diagonal or angular, it can always be reduced to a known constant. Then should do well to apply the ordinary textbook methods of least squa successively to each equation of the system. If the disturbance variab in the several equations are mutually independent and at the same ti the Jacobian expression is triangular (a recursive system), then a p ticular arrangement of least squares estimates is identical with ma mum likelihood estimation. It should be obvious that complicati arise if the disturbances for different subscript values (say, differ time-periods) are not mutually independent. Likewise, the assumpt that the disturbances are normally distributed is convenient but necessarily essential.

All these issues are brought out neatly in the well-known cobw theory applicable to demand and supply in many agricultural mark Let P_t = price, D_t = demand, S_t = supply, and u_t, v_t, w_t = rand disturbances; other symbols are constants. We formulate a dema supply, and market-clearing set of equations

$$P_t = \alpha_0 + \alpha_1 D_t + u_t$$
$$S_t = \beta_0 + \beta_1 P_{t-1} + v_t$$
$$S_t = D_t + w_t$$

By substitution we reduce this to

$$P_t = \alpha_0 + \alpha_1 D_t + u_t$$
$$D_t = \beta_0 + \beta_1 P_{t-1} + x_t$$

where $x_t = v_t - w_t$.

With u_t and x_t independent of each other, we find their joint norm distribution over the sample range to be

$$\left(\frac{1}{2\pi\sigma_u\sigma_x}\right)^T \exp\left\{-\frac{1}{2\sigma_u{}^2}\sum_{t=1}^{T} u_t{}^2 - \frac{1}{2\sigma_x{}^2}\sum_{t=1}^{T} x_t{}^2\right\}$$

ubstitution of expressions for x_t and u_t, we get

$$\begin{vmatrix} \alpha_1 \\ 1 \end{vmatrix}^T \left(\frac{1}{2\pi\sigma_u\sigma_x}\right)^T \exp\left\{-\frac{1}{2\sigma_u{}^2}\sum_{t=1}^{T}(P_t - \alpha_0 - \alpha_1 D_t)^2 - \frac{1}{2\sigma_x{}^2}\sum_{t=1}^{T}(D_t - \beta_0 - \beta_1 P_{t-1})^2\right\}$$

ıe maximisation of this joint distribution it is necessary to minimise

$$\sum_{t=1}^{T}(P_t - \alpha_0 - \alpha_1 D_t)^2 \text{ with respect to } \alpha_0 \text{ and } \alpha_1$$

$$\sum_{t=1}^{T}(D_t - \beta_0 - \beta_1 P_{t-1})^2 \text{ with respect to } \beta_0 \text{ and } \beta_1.$$

the earlier days of econometrics, investigators were always in ɔt about the direction in which to form their regressions—quantity rice or price on quantity. In this example the procedure is un-iguous: form the regression of quantity on lagged price to obtain upply equation and the regression of price on quantity to obtain demand equation. By writing the system in this way we have ded all problems of the Jacobian expression, which takes the value /.

the problem were not as simplified as in our cobweb example with pendent normal disturbances, maximum likelihood estimation ıt be more difficult. This method has many nice optimal pro-es; therefore we like to use it or set it up as a norm against which ıdge alternative estimates. To go into the difficulties connected maximum likelihood estimation for equation systems, and the ıical compromises made in order to render the problems more ageable, would lead us far astray in this general scanning of the of econometrics. We merely observe here that the problem of ıation can be attacked with the artillery of mathematical statistics, various approaches can be developed. At a simple level, many istic or intuitive methods can be devised. In addition to the ıod of least squares there is the method of instrumental variables, ɛcial case of which is the technique known as subgroup averages. riant of maximum likelihood is frequently used in cases in which above cobweb-type simplifications are not applicable. Least res methods are familiar to all, but in the present context they to biased estimates. The other methods mentioned, if applied erly, are statistically consistent.

ıother problem plaguing the earlier workers in econometrics has ɛ to be known as the identification problem in the modern oach. This has to do with the theoretical constructs of economics. near models we may have no way of distinguishing some par-ar equation of a system from others, or linear combinations of . In the cobweb example we reduced the system to two linear tions in quantity demanded and price. Fortunately price is lagged

in one equation but not in the other. In this way we achieve identifi tion of a precise type. Had all variables the same subscript (no la we should have been faced, as statisticians, with two linear equati in the same variables, with additive disturbances, and unknown pa meters. From the point of view of the statistician there is no way distinguishing between them. Situations like these gave rise at an ea stage of econometrics to recognition of the fact that one could determine both a supply and a demand equation simultaneously fr the same set of data on prices and quantities alone. The problem identification is still with us in a very practical sense, but now understand it better; in fact, full conditions for identifiability in lin systems can be conveniently formulated. In an abstract form they c also be stated for more general systems.

Problems Studied by Econometrics

The formulation of the typical problem outlined in the preced section applies to a wide variety of specific situations, i.e. to any p of the economic process that can be described by a system of equatio The reference to the cobweb model of supply and demand is an il stration. This is a particular and important field of application.

The earliest contributions to econometrics were attempts to establ statistical demand functions. Economists had long been drawing sweeping downward stroke to describe a demand function to studer without any concrete evidence about the existence of such functi or their main properties. At most, the pure theorist can distingu between the price and income elasticities of necessities and luxur the former type of goods being broadly considered to have lower e ticities than the latter. From the pioneer studies of Henry Schu down through the more recent studies of Fox, Stone, and Wold have had systematic statistical studies of the demand for many fo and other agricultural products. These have been mainly dema studies, but a few supply relations have also been estimated. Ma workers have contributed to a vast literature on statistical dema functions covering the products of several countries and extending other items in consumers' budgets.

A great part of this work has been guided by the successively refir developments in the pure theory of consumer behaviour. Investigat have paid attention to the homogeneity properties of demand functio implied by classical economic analysis and the interrelations amo cross-elasticities associated with the mathematical conditions extrema of consumers' utility. The classification of goods, dependi on the degree of elasticity, has been made possible by these studies

The supply side of markets may be approached formally throu what economists call supply functions or through underlying and mo basic relationships—the production and cost functions. We associa with the name of Douglas the studies of aggregate production fur tions, linear in the logarithms of inputs and output. A basic objecti

these studies has been the testing for existence of decreasing, con-
ant, or increasing returns to scale and the applicability of the marginal
oductivity theory of wage remuneration. Douglas's world-wide
udies of manufacturing production have led him to conclude that
ere is much evidence in support of constant returns to scale and the
ayment of labour's 'just' reward according to the marginal produc-
ity theory of wages. These same conclusions have not been reached
all industries.

The analysis of statistical cost functions, associated to a large extent
ith the name of Joel Dean, has pursued another basic issue of eco-
omic analysis, namely the linearity of the total cost function or, in
her words, the constancy of marginal cost. Statistical studies of cost
nctions mainly concluded that total cost functions are approximately
ear.

Statistical studies of foreign trade equations may be covered under
emand analysis. The home demand for imports and foreign demand
r exports are actually particular cases of demand functions, but such
great effort has been devoted to the statistical determination of
port-import elasticities because of their importance in understanding
e economic mechanism in a large part of the world that these econo-
etric investigations deserve special mention.

Much the grandest and most sweeping use of econometric methods
curs in the estimation of business cycle models of the economy as a
hole. We try to describe the workings of an entire economy by a
stem of simultaneous equations. These are highly aggregated rela-
onships and contain the summation of detailed demand functions in
hich investigators of demand behaviour, as such, are interested; the
obal production function of the Douglas type; aggregate export-
port equations; the demand for fixed and working capital; the for-
ation of prices; the distribution of national income between wages
d profits; equations of the money markets; and similar relationships.
ch econometric systems are often associated with models of the trade
cle. They have in fact been estimated for the USA, Holland,
nada, and other countries.

The great bulk of what is commonly known as econometric research
s been concerned with the problem of estimating economic relation-
ips, as outlined above. This is one of the more controversial aspects
econometrics, and perhaps the most interesting, yet there are other
oblems which are either clearly different or only indirectly related
this central matter. In order to give the reader a truer impression
the scope of econometrics, these problems are briefly mentioned.
st, we note the problems of distribution, i.e. distribution in the tech-
al statistical sense. Econometricians have long been fascinated by
e empirical regularities of the distributions of income, wealth, and
ilar economic variables, yet have not derived ultimate mathe-
tical laws that would be nearly universally applicable. The prob-
n has been to find distributions that are more satisfactory than the

extremely suggestive Pareto and logarithmic normal distributions income and wealth. A more basic problem in this connection has been tackled, namely to develop stochastic processes explaining generation of these distributions from elementary considerations ab the way people move about between income and wealth classes.

The problems of distribution are, of course, not unrelated to central problem of estimating relationships, since they are often cru in building a bridge between micro-economic and macro-econo relations. A more general problem in this connection, however, is aggregation or index number problem. To some extent, index numb are purely the subject matter of descriptive economic statistics; there are important analytical problems of an econometric natu since our underlying theories are usually about individual people a individual commodities, whereas our data usually refer to ent markets and large collections of people.

In the theory of production economics and the theory of the fi the subjects of linear programming and input-output analysis cont econometric offshoots. When these two approaches come to empirical stage of gathering data and estimating coefficients or ent relations, they involve econometric problems. The empirical work these fields has not been properly tied in with probability theory a statistical inference, yet that is not because the problems are outs the bounds of these two subjects. It is simply because they have yet been formulated properly as econometric problems, as has the ot work in estimating economic relations, and because the empirical d used have not constituted a reasonable probability sample.

Applications of Econometrics

The pure theory of econometrics, the theory of how best to meas in economics, can stand by itself as an objective, but we might v query the uses to which applied econometrics can be put. Ecc metric work could be given a blanket justification as an attemp give empirical content to theoretical economics. It is not sim measurement, in itself, that gives this empirical content to theory is only measurement of the relevant variables followed up by attempt to put the variables together in the patterns suggested hypothesised by theory. Lecture-room pedagogues ought to be a to tell their students about some properties of actual demand functio production functions, import propensities, etc. that they talk about the abstract. Econometrics serves the useful purpose of enabl theoreticians to clothe their concepts in concrete quantitative ter With some confidence they can speak of the demand for food in ma countries as inelastic with respect to both income and price variatio they can similarly describe as more elastic the demand for many c sumer durables that have been investigated by econometricians. both cases, they can give approximate numerical magnitudes. mentioned in the previous section, theorists on the behaviour

terprise can speak of the cases of constant returns to scale (production nctions) and constancy of marginal cost (linear total cost functions) the basis of substantial econometric work. Similarly, we have a r idea of the main spending propensities, multipliers, etc. derived om actual econometric models of the economy as a whole for use in scribing the workings of business cycle theory.

A stage beyond the mere numerical description of theoretical con-ucts is the testing stage. Much of economic theory is based on very ecial assumptions about human behaviour, and it behoves us to look to the validity or reasonableness of these assumptions. At the risk of ing too repetitive of what has been used as an example on many other casions, we might mention that econometric evidence has shown tle support for the simplest statement of the acceleration principle of vestment, a theory which, if true, would greatly facilitate develop-ents in business cycle theory. People who are devoted to the approach the acceleration principle have been forced to modify and improve is theory of investment to make it more acceptable. In the same way appears that much of the widespread testing of Keynes's ideas about e propensity to consume will show that the simplest formulation is adequate; and that a better, more elaborate theory will eventually forthcoming on the basis of current econometric studies.

To some extent, we are able to reject hypotheses, but our samples e too limited to enable us to prove that certain theories ought to be cepted to the exclusion of all competitors. The usual situation is at many competing theories will explain equally well the same body manifest data. By continually testing and sifting, we hope to narrow e acceptable range of hypotheses, but it is surely too much to expect ore from econometrics (or from any other method for that matter) the near future.

The ability to predict is often said to provide a strict test of a theory. extending our testing grounds to the field of prediction, we greatly rich the chances of rejecting faulty theories, because in observed mples we tend to force a fit or agreement between theory and actice. The real test of an empirical result, then, is its functioning tside the sanctuary of the sample, where data are free to vary. No onomic theory is complete; there must always be some exogenous riables; hence the predictive type of test must appropriately dis-guish among discrepancies in the endogenous mechanism, dis-epancies in the exogenous factors (pre- and post-prediction periods) d chance disturbances. Prediction, then, is understood in a broad ıse—as scrutiny of the econometric results outside the confines of e sample to which the equations are fitted.

In another sense, however, prediction serves as an application of onometrics. We want prediction in its own right as an objective gardless of its function as a test of theory, and a very practical plication of econometrics is its use in forecasting things to come. onometrics has by no means solved the inherently difficult problem

of forecasting, nor has any other method, but it does, in my opini make a strong contribution in this area. It serves as a systematic te nique for gathering together historical facts, in so far as they statistical or quantitative, and interpreting them in a way that he us to judge the future on the basis of the past.

A number of interesting theoretical problems arise in connecti with forecasting. First of all we must devise criteria for classifyin forecast as good or bad and for deciding whether an estimated the has performed within limits set by sampling error. Problems of ser correlation of errors also present interesting situations. It would see that promising theoretical contributions remain to be made on t entire forecasting problem along the lines laid out by the theory statistical decisions.

It should be recognised by anyone who has attempted forecasti that by no means all relevant information is statistical or contain in the main series that have been compiled over a number of yea An interesting issue in the econometrics of forecasting is how to ble last-minute developments, results of sample surveys, and various b of non-quantitative information available to the forecaster, with t mathematical-statistical model being used.

Realistic and true forecasts have frequently been made from agg gative models of the economy as a whole. These are not purely endo nous models, but they are designed so that the forecaster need ass *a priori* values only to variables about whose future course we f reasonably confident. The estimation of partial models is widespr in econometrics (e.g. demand, supply, production, and cost functio but these generally serve as tools for conditional forecasts. To so extent, applications in agricultural economics, which have been wi spread, have met with a measure of success in price forecasting. Th predictions are as independent of unknown conditions as are forec for the general economy in the aggregate.

Major Outstanding Problems and Limitations

We econometricians like to feel that we have made rapid intellect progress in the last two or three decades. We certainly have be materials to work with and supposedly sharper tools, but serious pr lems remain. These are essentially the same problems that confron the earliest workers in the field, but our main advantage today probably that we understand these problems more clearly and c formulate them more sharply.

One of the first problems to be recognised, that of collinearity amo the several variables of any particular equation being estimated, still with us. Economic variables move roughly together, and t nature of our non-experimental data is such that we cannot, except extreme good fortune in some circumstances, gather together a sam that shows the independent variation of the strategic effects we seeking. In export-import equations, income and terms-of-tra

riables move together so that we have difficulty in attributing observed enomena to income effects or terms-of-trade effects. This remark rries over to demand analysis in general, except for the possibility at estimates of total (foreign and domestic goods) demand elasticities n be partially determined from consumer survey data (income sticities) and partially determined from market statistics over time rice elasticities). The pooling of micro-economic cross-section data d macro-economic time-series data, while raising formidable fresh oblems, does undoubtedly help to overcome the collinearity problem. his problem has been equally important in separating marginal contributions of capital and labour in production functions, and on a der scale in separating the effects of a large number of variables in acro-economic models of the trade cycle.

Three important problems confronting the theoretical statistician ho is devising methods of estimation and hypothesis-testing in econoetrics are (i) treating non-linear systems; (ii) dealing with serially rrelated disturbances; and (iii) dealing with errors of observation. The principles of statistical inference used in econometrics are veloped on a general basis only for linear equations, either singly in a system. The statistician's unknowns are not the economic riables entering the equations; they are the parameters. Strictly eaking the models to be estimated or tested need only be linear in e unknown parameters, thus leaving a good deal of, but not unlimited, ope for the introduction of non-linear economics. In so far as econosts are interested in making statements about the dynamic properties of a system, they can make statements of general applicability ly for linear systems. For purposes of prediction linear systems are mparatively simple to solve for the predicted values in terms of the edictors. Non-linear systems can be solved approximately for ecific forecast values, but often this involves more work than with a ear system of equivalent size. As for the general dynamic properties the system, however, there is little that can be said for a wide class non-linearity. Non-linear systems can assume an infinite variety of ms, and we can assess the dynamic properties only for specified oes of non-linearities. Possibly with further developments on the athematical side, we shall be able to widen the class of non-linear tems for which general statements can be made. Moreover, econosts may be able to specify more precisely the types of non-linearities be encountered. Empirical econometric research should also be of lp in this respect.

Another respect in which linearity has proved to be an extremely nvenient simplification is in the formulation of conditions for identifility. Simple counting rules and tests on the rank of certain matrices l us whether or not equations in a linear system can be identified. e criteria can be formulated for the general case, including all forms non-linearity, but it is no simple matter to determine whether parular non-linear models are identifiable.

A hopeful sign in the analysis of non-linear systems is the wider of high-speed electronic computing devices. For example, a mach of large capacity (in today's sense of the term) can quickly g approximate numerical solutions to a non-linear system which wo otherwise require an inordinate amount of hand labour. The mach operation can be quickly repeated over and over again; hence arbitrary assumptions about the course of exogenous variables time paths of endogenous variables can be numerically computed a plotted. In this way the dynamic properties can be examined. I similar fashion various types of analogue computing machines can a be used to study non-linear dynamics.

For the moment, we frequently try to approximate non-linear s tems by piecing together linear segments in much the same way as arc can be geometrically approximated by sections of a polygon. T approach usually requires the resort to larger linear systems th would otherwise have been required and will prove to be a good su stitute for genuine non-linearities only if we are able to accommod the larger computational problems involved. Again, we seek relief the use of high-speed, large-capacity computing machines.

The second front on which mathematical statisticians can contrib effectively to the development of econometric technique is in the tre ment of serially correlated random disturbances. This problem under attack now, and developments in recent years have contribu towards the devising of tests to detect serial correlation and of meth of allowing for its influence in the estimation of parameters of di economic interest. Nevertheless, problems remain. Linear syst with linear lagged correlations among disturbances can be handl but the computational burden is increased to a large extent. M general forms of lack of mutual independence among disturbar remain to be tackled. The problem of serial correlation is of obvi importance in forecasting, and some comparatively simple pract measures can be devised to meet this issue, but a systematic theoret treatment would be most helpful.

Some econometricians distinguish between 'shock' and 'er models, i.e. between systems in which the probabilistic elements a as disturbances in entire equations, say through omitted explicit v ables, and systems in which these elements arise as errors in the in vidual variables, say as observational errors. The typical probl of econometrics as formulated above (pp. 1–6) may be called typical 'shock' model, and it has been given a fairly full and gene treatment. The 'error' model is closely akin to the theory guiding pioneers of econometrics, and it has been found to lead to meaning estimates of parameters only if *a priori* restrictions are placed on variance-covariance matrix of errors. In most cases this proves to an inherent difficulty.

The choice between the two types of stochastic model is unfortun since both types of random element exist in fact. We are far fr

ving included all the relevant variables explicitly in our models and e far from having perfected measurement techniques at the stage of ta collection; therefore a fully general treatment of methods to be ed in shock-error systems would be welcome. Some results along is line have been obtained, but basically an outstanding problem mains.

Economists have exposed themselves to the accusation of being called zy' by drawing sharp lines demarcating the limits of their subject to ake it distinct from other fields in which human behaviour is studied. his affects econometric work in two significant respects. In the first ace it has led to the classification of some variables as exogenous hen they really ought not to be so classified, and in the second place has led to the inclusion of some possibly important factors in the sidual random disturbances. In practice government variables (taxes, ansfers, expenditures); foreign variables (incomes and prices abroad); mographic variables (population, birth rates, immigration, age, mily size); climatic variables (temperature, rainfall, sunshine); and riables external to particular markets being studied are placed in e exogenous category. To a large extent, this choice is indicative of willingness or temporary inability to develop a theory embracing ese variables as endogenous. It is only for climatic variables that we n make a *pure* case for inclusion in the exogenous category. The hers are affected to a greater or smaller extent by the very economic ocess being studied. The problem remains of widening current onometric practice by including more of the so-called exogenous riables in the endogenous category and treating them correspondingly the stage of statistical estimation. The long-run growth models now ing developed have broken through these restrictive practices by eating population and associated demographic variables as endo- nous. This procedure deserves enlargement on many fronts.

The second respect in which the limits of economics have been too rrow is in the neglect of various subjective factors. They have been egated to the 'catch-all' disturbances as being part of 'psychology' other non-economic disciplines. A stage prior to the incorporation subjective magnitudes in endogenous models is the primary stage measurement of subjective magnitudes and the use of them as pre- termined variables in behaviour equations. The most promising actical applications now appear to come from field surveys of indi- lual economic units in which questionnaire information on their itudes, beliefs, expectations, intentions, etc. is collected. Economists ve often talked about expectations and their role in behaviour; in t problems are often reduced to a range of alternative possibilities pending on the state of expectations. We are now in a position, tunately, to attempt actual measurement of expectations and other bjective magnitudes on the basis of personal interview surveys. With- t taking up the difficult problem of integrating these magnitudes o a determinate system as endogenous variables, we can at least try

to ascertain the impact of these variables on economic decisions, ticularly by lagging them, i.e. by studying the relationship betweer subjective variables on a prior date (first interview) and the ac variables, such as spending or saving, on a later date (second or sequent interviews). This type of econometric analysis coupled a study on the temporal stability of individual attitudes will be of in decomposing the random disturbances, understanding behavi and making predictions. The instruments of measurement used in respect are notoriously crude; therefore much remains to be done the interesting problem of *scaling* subjective variables that are us in the construction of our econometric models. Possibly much can borrowed from psychological theories of scaling in this respect.

Sample survey data have much greater uses in econometrics bey the problem just mentioned of ascertaining subjective variables. D from this source have long been used in demand analysis, the estimat of production functions, and other applications. Modern surveys usually made with an efficiently designed representative sample, these designs are complex in comparison with the simple and p random sample put up as a norm. Many econometric studies based these data have conveniently assumed that the underlying sam design was of the simple random type, but an interesting proble posed for statisticians to gear econometric and other derived compu tions more closely to the sample design that is, in fact, used.

When a sampling statistician draws a probability sample for a f survey, he nearly always has a firm idea of the universe being samp It may be all the family or consumer units in a specified geograph area, it may be all the firms in an industry, it may be all home owr etc. In so far as econometric research is based on this type of d sample and universe are well defined. The samples are also freque large, containing hundreds or thousands of items. When using o samples of data, particularly time-series samples, the econometri is not so fortunate. His sample then is, comparatively speaking, sr frequently consisting of only twenty to thirty observations. Equ serious is the problem of relating this sample to a well-defined univ It is seldom a properly selected sample; it consists simply of the f ments that can be pieced together, and one is grateful for every s of evidence that can be included. It is not a representative sampl the universe of historical time-periods, and one should be wary drawing general conclusions from a sample that does not contai wide variety of economic environments. In fact, wartime years usually excluded from analysis, being considered as abnormal abe tions. One econometrician has jokingly put it this way: 'We d have enough wars to provide a reliable sample of wartime observa periods.' The way out of the difficulty is suggested by assuming universe to be the infinitely large sample of all possible econo decisions that could have been made in the historical environm Our sample consists of the few decisions actually made from

initely large collection. We assume, without firm basis, that a nple so obtained is a random sample. When possible, we recognise e customary smallness of the sample and apply small sample theory, t in many circumstances we use large sample theory for our handful observations. In my opinion, these sampling difficulties clearly pose limitations on econometrics today.

Following in a similar vein, it should be pointed out that econo-tricians often assume that the random disturbances follow the normal obability law. Not only is large sample theory widely used, but also rmal theory is widely used. Some procedures are non-parametric general for a wide class of distribution functions, but much of the rk rests on an assumption of normality. Compared to the other oblems, this may be of a much smaller order of magnitude. In gregative systems summed over people or over commodities, nor-ality assumptions may be reasonable and in some cases statistical ocedures have actually been extended to non-normal populations. more serious difficulty than the error in specification of the form of e probability distribution is the error in specification of the entire tem. The econometrician draws upon the best *a priori* information ailable, but that may be inadequate for the enormous task of setting t a list of the relevant variables involved and the form of the equa-ns in which these variables are interrelated. Specification errors are be reckoned with, as are the other problems mentioned in this tion. The result is that no one knows better than the econometrician e limitations of his subject. His main comfort is that he realises that is trying to answer intrinsically difficult questions of economics and certainly no worse off than those using alternative tools.

SELECTIVE BIBLIOGRAPHY

eral

AVELMO, T. (1944). 'The probability approach in Econometrics', *Econometrica*, **12**, Supplement.

OD, WM. C., and KOOPMANS, T. C., Editors (1953). *Studies in Econometric Method*. Wiley, New York.

EIN, L. R. (1953). *A Textbook of Econometrics*. Row, Peterson and Co., Evanston.

OPMANS, T. C., Editor (1950). *Statistical Inference in Dynamic Economic Models*. Wiley, New York.

BERGEN, J. (1951). *Econometrics*. The Blakiston Co., Philadelphia.

TNER, G. (1952). *Econometrics*. Wiley, New York.

nand Analysis

LEN, R. G. D., and BOWLEY, A. L. (1935). *Family Expenditure*. Staples, London.

K, K. (1953). *The Analysis of Demand for Farm Products*. Technical Bulletin No. 1081. US Department of Agriculture.

UTHAKKER, H. S., and PRAIS, S. J. (1955). *The Analysis of Family Budgets*. Cambridge University Press.

ULTZ, H. (1938). *The Theory and Measurement of Demand*. The University of Chicago Press.

STONE, R., and Associates (1954). *The Measurement of Consumers' Expenditure Behaviour in the United Kingdom* 1920–1938. Cambridge University Pr
TOBIN, J. (1950). 'A statistical demand function for food in the USA', *J. R. St Soc.*, A, **113**, 113.
WOLD, H., and JURÉEN, L. (1951). *Demand Analysis*. Wiley, New York.

Production and Cost Functions

DEAN, J. (1936). *Statistical Determination of Costs with Special Reference to Marg Costs*. University of Chicago Press.
DOUGLAS, P. H. (1948). 'Are there laws of production?', *American Economic Rev* **38**, 1.
MARSCHAK, J., and ANDREWS, W. H. (1944). 'Random simultaneous equations the theory of production', *Econometrica*, **12**, 143.

Export-Import Functions

CHANG, T. C. (1951). *Cyclical Movements in the Balance of Payments*. Cambridge U versity Press.
MORGAN, D. J., and CORLETT, W. J. (1951). 'The influence of price in internatio trade: a study in method', *J. R. Statist. Soc.*, A, **114**, 307.
NEISSER, H., and MODIGLIANI, F. (1953). *National Incomes and International Tr* University of Illinois Press, Urbana.
POLAK, J. J. (1954). *An International Economic System*. University of Chicago Pre

Aggregative Models

BARGER, H., and KLEIN, L. R. (1954). 'A quarterly model for the US econor *J. Amer. Statist. Ass.*, **49**, 413.
CLARK, C. (1949). 'A system of equations explaining the United States trade cy 1921 to 1941', *Econometrica*, **17**, 93.
KLEIN, L. R. (1950). *Economic Fluctuations in the United States*, 1921–1941. W New York.
KLEIN, L. R., and GOLDBERGER, A. S. (1955). *An Econometric Model of the U States*, 1929–1952. North Holland Publishing Co., Amsterdam.
TINBERGEN, J. (1939). *Statistical Testing of Business Cycle Theories, Vol.* 2, *Bus Cycles in the United States of America*, 1919–1932. League of Nations, Gen

Sample Survey Information

DUESENBERRY, J. (1949). *Income, Saving and the Theory of Consumer Behaviour*. Har University Press, Cambridge.
KLEIN, L. R., Editor (1954). *Contributions of Survey Methods to Economics*. Colun University Press, New York.
MENDERSHAUSEN, H. (1940). 'Differences in family savings between cities of diffe size and location, whites and negroes', *Review of Economic Statistics*, **22**,
NICHOLSON, J. L. (1949). 'Variations in working class family expenditure', *J. Statist. Soc.*, A, **112**, 359.
TOBIN, J. (1951). 'Relative income, absolute income and saving', *Money Trade Economic Growth*, pp. 135–56. Macmillan, New York.

Miscellaneous

AITCHISON, J., and BROWN, J. A. C. (1954). 'On criteria for descriptions of inco distribution', *Metroeconomica*, **6**, 88.
ANDERSON, T. W., and RUBIN, H. (1949–50). 'Estimation of the parameters o single stochastic difference equation in a complete system', and 'Asympto properties of estimates of the parameters of a single equation in a comp system of stochastic equations', *Ann. Math. Statist.*, **20**, 46; **21**, 570.

AMPERNOWNE, D. G. (1953). 'A model of income distribution', *Econ. J.*, **63**, 318.
BIN, J., and WATSON, G. (1950–51). 'Testing for serial correlation in least squares regression', *Biometrika*, **37**, 409; **38**, 159.
IMURA, S. (1954). 'Toward a general non-linear macrodynamic theory of economic fluctuations', *Post-Keynesian Economics*, pp. 192–226. Rutgers University Press, New Brunswick.
NSTON, J. (1955). 'Econometric models and the average duration of business cycles', *Manchester School*, **23**, 193.
CUTT, G., and COCHRANE, D. (1949). 'Application of least squares regression to relationships containing autocorrelated error terms', and 'A sampling study of the merits of autoregressive and reduced form transformations in regression analysis', *J. Amer. Statist. Ass.*, **44**, 32 and 356.
EIL, H. (1954). *Linear Aggregation of Economic Relations*. North Holland Publishing Co., Amsterdam.

THE EFFECT OF OBTAINING INFORMATION FROM DIFFERENT INFORMANTS ON A FAMILY MORBIDITY INQUIRY

ANN CARTWRIGHT
London School of Hygiene and Tropical Medicine

In survey work, information about the various members family can be obtained by different methods of interv Miss Cartwright here describes an investigation designe study the effect of these different methods on the informa obtained in a certain type of inquiry.

The results given here were obtained in a small pilot inquiry wh was carried out by the Public Health Department of the London Sch of Hygiene and Tropical Medicine. This department is engaged o programme of research, the broad aim of which is to study people's of the different parts of the National Health Service. The whole stu is concentrated on a relatively small area, so that material can collected from several sources about the different aspects of the probl A family morbidity survey is designed as a part of this research proj to supplement information obtained from various records, and present a picture of the problems of ill-health and the success of health services in solving these problems, as seen from the viewpoin the individuals and families concerned.

The main aim at the first pilot interview for the family morbi inquiry was to ask about the health of all the members of a sampl families and to find out about the progress and treatment of all illne and disabilities which were reported. At two subsequent intervi which took place at four-weekly intervals, information was sou about the illnesses, consultations, and other events related to he that had occurred since the previous interview.

One of the first decisions that had to be made was: which mem or members of the family should be interviewed. The suggestion t information concerning all members of the family should be obtai from a single informant, the housewife, had the attraction of simpli and economy; there were, however, grounds for suspecting that s a procedure might have serious disadvantages.

It is extremely difficult to define health or illness objectively. If treatment of the illness is taken as the criterion, this can be defi fairly easily; but this would have been inappropriate in an inqu designed to relate the use of services to personal assessment of illn Any statement or assessment of illness is likely to be subjective therefore to depend on the attitude of the informant. It seems lik

t a person will not regard his or her own ill-health and the ill-health ɔther people in quite the same way, and that the assessment of ther person's ill-health will depend upon his relationship with the ɔrmant. Another factor which might affect the level of reporting is completeness and accuracy of the informant's report. A housewife h a large family might become bored and inattentive or forgetful if interview lasted too long. Finally, the housewife might not possess the information required. Because of these considerations it was ided to make an attempt, on the pilot inquiry, to discover whether housewife was an adequate informant for all the adults in the ily.

The pilot sample consisted of ninety households and was confined to small area of the estate where the main inquiry was to be held. ɔne-in-four sample of households was selected systematically in lress order and the selected addresses were allocated in rotation to ee interviewers. The addresses on each interviewer's list were then ɔcated alternately to two groups. In the first group the interviewers e instructed to interview all persons in the household aged sixteen nore personally, and to obtain from the housewife information about self, about the health of children under sixteen, as well as certain iological details about the family; but not about the health of other ılts. In the second group they were instructed to interview only the ısewife and to obtain from her information about all the members the household. It should be emphasised that the design of this uiry does not make possible a strict comparison between information ained solely from the housewife about the entire family and the ɔrmation provided by all the adult members of the family. The ısewife, even in those families where she was intended to be the sole ɔrmant, could not always be interviewed in privacy. There were a instances in which other adults were present at the interview, and ectly or indirectly made unsolicited contributions to the information lected. The comparison which follows is between the two methods nterviewing.

e Response Rate in the Two Groups

Only two-thirds of the pilot sample of families co-operated com-tely for the three interviews. This is not surprising when it is nembered that for half the sample all the adults had to be inter-wed on three separate occasions, and that failure to contact any ılt member of the family at any one time meant that the family was fully covered. The numbers involved are small, but they suggest ather high failure rate at the first interview for the group in which housewife only was to have been interviewed. However, once a t interview had been obtained there was relatively little loss at sequent interviews in this group. In the group where all the adults l to be interviewed rather more families co-operated at the first erview, but after that the response fell more quickly, and finally the

proportion of families completely covered in the inquiry was sm for this group than for the other.

These differences in response rates, which are shown in Table I not indicate any very significant difference in the sample achieved the two groups.

TABLE I

The response rate in the two groups

	Method A: All Adults Interviewed Personally		Method B: Housewife only Interviewed		All* the Sample	
Families allocated	43 ≡ 100%		44 ≡ 100%		90 ≡ 100	
Successfully completed:						
At first interview	39	91%	35	80%	76	84
At second interview ..	32	74%	33	75%	67	74
At third interview	28	65%	32	73%	62	69

* Three families, known to have only one adult, were not allocated to either group.

The Number of Illnesses Reported

Three questions were asked about illnesses at the first interview. T first one related to disabilities and chronic illnesses; the second to ot diseases, injuries or minor complaints present at the time of intervi while the third question consisted of a list of diseases and parts of anatomy similar to the one used by the Government Social Survey the Survey of Sickness.

> 'Have you got anything wrong in the way of colds, catarrh or t and throat trouble or anything wrong with your eyes, ears, te head pains, chest, heart, stomach or indigestion, liver, kidn bowels or constipation, legs, feet, hands, arms or rheumatism, s complaints, infectious diseases, or anything wrong with your ner To All Women. Have you anything wrong in the way of wom complaints?'

Interviewers were told that, when they were asking about more t one person, they should ask the first two of these questions for e member of the family in turn, and after that go on to ask the th question once about all the people concerned. This was becaus had been found impractical, on a small pre-pilot study, to repeat check list in the third question.

As soon as the informant mentioned any complaint, the intervie recorded the description of the illness and then asked either 'How d it affect you (or him/her)?' or 'What do you think is the cause?' T interviewer then continued with the questions to discover what ot illnesses were present. She did not go on to ask the other questi about each complaint until all the illnesses had been elicited. T order was adopted because the informant might have been reluct to mention further illnesses if he had known how many questions

uld be asked about each complaint. There may still have been an irect effect of this kind through the interviewer, who was well aware he large number of dependent questions.

To compare the effect of the two methods of interviewing on the orting of illness it is simplest to look at the information recorded out married male tenants and their wives. This reduces the size of sample only slightly, since 36 out of 39 families on Method A and out of 35 families on Method B contained a married male tenant ng with his wife.

Table II shows the number of illnesses reported by, or for, married le tenants and their wives in the two groups.

TABLE II

Number of illnesses reported at first interview

	Married Male Tenants				Wives of Tenants			
	Method A: Interviewed Personally		Method B: Information Obtained from Wife		Method A: Interviewed about Self and Children		Method B: Interviewed about All Household	
Number of people	36		33		36		33	
Number of illnesses reported	91		27		62		43	
Number of illnesses per person	2·5		0·8		1·7		1·3	
Distribution of illnesses per erson:	No.	%	No.	%	No.	%	No.	%
0	3	8	15	49	8	22	9	27
1	9	25	13	36	11	31	13	40
2	10	28	2	6	8	22	5	15
3	3	8	2	6	3	8	4	12
4	5	14	1	3	5	14	2	6
5	4	11			0	0		
6	1	6			1	3		
7	1							

The results for the married male tenants are surprising. The group t was interviewed personally reported an average of two and a half esses, while in the other group, where the information was obtained n the wives, less than one illness per man was recorded. When the ribution of illnesses is considered, less than 1 in 10 of the married n interviewed personally said they had no illness, but in the other up nearly half the wives reported that their husband had no illness. The number of illnesses recorded for the wives does not differ nearly much between the two groups. However, slightly more illnesses e reported by the wives who only had to answer questions about mselves and the children under sixteen than by those who also had nswer for the other adults in the household. The difference is not istically significant on this small sample; but it may nevertheless be eal difference, for the amount of information an individual gives

*

about his own health seems definitely to be influenced by the perso range of his report—that is to say, whether or not he deals with health of other people as well as with his own. This is indicated by comparison of married male tenants with their wives in the grou which all adults were interviewed. Here, more illnesses were repo by the men than by their wives. This is contrary to what would be pected from other studies of morbidity, and the likely explanation is t the men were answering questions only about themselves, whereas th wives had in addition to give information about the children's hea

At the second interview, which took place approximately four we later, the interviewer went over the illness recorded at the first in view and after that went on to ask three questions to discover whet the person had had any other illness or complaint since her previ visit. [The three questions asked first (question 4) about any illness; th (question 5) about any accident or injury; and finally question which corresponded to the check list at question 3 on the first int view.] After this, the person was asked for details of hospital vis GP consultations, other medical or health consultations, any incapaci and medicines taken since the first interview. When anything w mentioned, one of the dependent questions was 'What illness was th for?' and if it was for a complaint not already recorded, the int viewer was instructed to go back and to record the appropriate det about the illness.

It might therefore have been expected that the group interviewed Method B, for whom fewer illnesses were reported at the first intervi would show more additional illnesses at the second. In fact, as sho in Table III, similar numbers of new illnesses were recorded at second interview for people interviewed by the two different metho This is so for both male tenants and their wives, and means that relative discrepancy between the two groups has been reduced.

TABLE III

Number of illnesses recorded at the first and second interviews

	Married Male Tenants				Wives of Tenants			
	Method A: Interviewed Personally		Method B: Information Obtained from Wife		Method A: Interviewed about Self and Children		Method B: Interview about A Household	
	No.	No. per Person	No.	No. per Person	No.	No. per Person	No.	No. per Per
Illnesses reported at:								
First interview	72	2·4	26	0·8	53	1·7	42	1·
Second interview ..	19	0·6	20	0·6	23	0·8	25	0·
Total	91	3·0	46	1·4	76	2·5	67	2·
Number of people ..	30		32		30		32	

Stages of the Interview at which Illnesses were Reported

t is interesting to look at the specific questions in answer to which esses were reported and to note the effect of the check list on the erent methods. For married male tenants, the use of a check list at first interview apparently increased the difference between the thods. However, even if the illnesses reported at the first two stions only are considered, the difference between the two methods ains substantial. For the wives of tenants the check list seems to e slightly reduced the difference between the two methods. These lts and those for the second interview are given in Table IV.

TABLE IV

The number of illnesses reported at different questions (number reported per person)

	Married Male Tenants		Wives of Tenants	
	Method A: Interviewed Personally	Method B: Information Obtained from Wife	Method A: Interviewed about Self and Children	Method B: Interviewed about All Household
irst interview:				
Q1. Disabilities and chronic illnesses.. ..	0·8	0·4	0·8	0·5
Q2. Other illnesses, injuries, etc.	0·2	0·2	0·1	0·1
Q3. Check list	1·3	0·2	0·7	0·7
Later	0·2	—	0·1	—
otal	2·5	0·8	1·7	1·3
econd interview:				
Q4. Any illness	0·1	0·3	0·5	0·4
Q5. Accident or injury ..	—	—	0·1	—
Q6. Check list	0·4	0·1	0·1	0·2
Later (at medicines, etc.)..	0·1	0·2	0·1	0·2
otal	0·6	0·6	0·8	0·8

he number of illnesses reported for the first time at the second rview is very small, but the proportion arising at the different stions indicates that the check list probably has the same effect e as at the first interview. Table IV suggests that questions about sultations and medicines slightly reduce the difference between the methods.

Type of Illness Reported

t the first interview, after all the illnesses had been recorded, the three of several questions asked about each illness were:

1. When did (you) first have it?
2. Have you ever been completely free from it since then?

If Yes to Question 2:

3. Are (you) troubled with it at the moment?

These questions make it possible to divide the illnesses into t categories: those that seem to the informant to be continuous; t that seem to be recurrent and are giving rise to symptoms at the of the interview; and those that are regarded by the informan liable to recur but are at present quiescent.

TABLE V

The type of illness reported by the two methods

	Married Male Tenants				Wives of Tenants			
	Method A: Interviewed Personally		Method B: Information Obtained from Wife		Method A: Interviewed about Self and Children		Method Interview about A Househo	
	No. per Person	%	No. per Person	%	No. per Person	%	No. per Person	
Continuous	0·7	28	0·4	48	0·5	29	0·5	3
Recurrent and troubling ..	0·4	16	0·1	19	0·4	23	0·2	1
Recurrent, not troubling ..	1·2	48	0·2	26	0·7	42	0·5	4
Not classified	0·2	8	0·1	7	0·1	6	0·1	
All illnesses	2·5	100	0·8	100	1·7	100	1·3	10
Number of people ..	36		33		36		33	

The greatest difference between the two groups of married m tenants is for illnesses which are recurrent but not troubling then the time of interview. In this category six times as many illnesses w recorded for the group interviewed personally as for those in which information was obtained from their wives. When only illnesses wh were present at the time of interview are considered, the differe between the groups is of the order of two to one.

The number of illnesses is too small to warrant a detailed classifi tion by disease. It is, however, interesting to note that whereas 10 of the 36 men interviewed personally said they suffered from ner none of the men in the other group were reported as having this trou

Conclusions

Although the size of the sample in this pilot inquiry was very sm there was a difference between the two methods of interview in number of illnesses reported for married men. How far this was d to the difference between obtaining personal information about man first hand or second hand, or to what extent it was because wife had to answer for several people, it is impossible to say. It see likely, however, from the evidence of the number of illnesses repor for the wife herself, that being asked about a number of people is lia to reduce the number of illnesses reported for each person. M research would be needed to ascertain the reason for the differe found here, and to find out anything definite about its nature.

There were two chief implications for the main family morbidity vey. It was clearly desirable to attempt to interview personally all adults in the families and thus to eliminate any bias caused by aining information at second hand. In addition it would obviously unsatisfactory to deal with the housewife's own health and that of children in the course of a single interview, as information about own health would not be strictly comparable to information given other adults who had been asked to provide it only about themves. It was thus necessary to have two sets of interviews with each usewife who had children, one dealing with her own health, at the ne time as the interviews with other adults in the family, and another, ew months before or after these, at which she was asked about her ldren.

nowledgements

The programme of research, of which the pilot study discussed in s paper formed part, was directed by Professor J. H. F. Brotherston, merly Reader in Public Health at the London School of Hygiene d Tropical Medicine, whose advice has been invaluable. I am teful also to Professor J. M. Mackintosh and other colleagues at the ndon School of Hygiene and Tropical Medicine for their helpful erest in this work.

LINEAR PROGRAMMING IN PRACTICE

EMIL KAY* and ERIC DUCKWORTH
The Glacier Metal Company Limited

Much has recently been heard of the technique known linear programming. In this article the authors show hov was applied to solve an industrial problem; they explain method of solution in broad outline, without giving all detailed computations.

Introduction

Linear programming is a technique used to solve simultaneo equations in which the number of unknowns exceeds the number equations. Without restrictions this circumstance gives rise to infinity of solutions. However, by defining that no solution shall negative and that a given linear function of the solution shall be maximum (or minimum), the problem can be solved, provided th such a solution does exist.

The linear function which is maximised is often a price or cost fu tion of the variables being evaluated, hence the ready application linear programming to economic problems. The application to described is one in which linear programming was used in determin the least-cost proportions of new and of recovered metal to be usec making up alloys to specification.

The computation will not be described in detail. The method u for arriving at the maximising solution was the so-called 'Simple due to G. B. Dantzig, which has been described elsewhere.[1, 2] T paper will describe the setting up of the matrices, the simplificati evolved, and the organisational consequences.

The Problem

Many plain bearings, of the type used in internal combust engines, consist of a steel shell on the inside of which is cast a th layer of whitemetal, composed mainly of tin with about ten per ce of antimony and four per cent of copper. Both the surface finish a the thickness of the whitemetal are carefully controlled so that a fi machining operation is necessary. The resulting scrap is collected a remelted and is known as 'recovered' metal: other sources of recover metal are scrap bearings from which the whitemetal is melted a surplus whitemetal from other ancillary operations. The recover metal is collected and cast into batches of ingots. Each batch is giv a bin number and stored until it can be used. It can rarely be us

* Now with NAAFI.

ectly for casting into bearings because of the alteration in composi-
n which occurs during handling and remelting.

The whitemetal foundry needs an almost daily requirement of white-
tal, and this is provided from a combination of fresh tin, antimony,
l copper, and of recovered metal in the correct proportions to give
exact composition of the metal needed when remelted together. As
re are fluctuations in the demand for fresh whitemetal and in the
pply of recovery melts, the daily requirements may vary from nil to
eral tons of different specifications and the number of recovery melts
ilable may vary from five to twenty-five. It is clearly desirable to
as much recovered metal as possible in making up fresh melts, but
usage is restricted by the fact that the metal required may differ
stantially in composition from the metal available and also that the
ximum amount of lead permitted in the whitemetal is 0·2%.

Recovered melts are classified into four categories according to their
d contamination:

A	less than 0·35%
B	0·35%–0·50%
C	0·50%–1·00%
D	1·00%–15·00%

Category D is due to the fact that lead-base whitemetals are some-
es produced in close proximity to tin-base ones and contamination
ery difficult to avoid.

With the figures for lead contamination in mind and with the
lisation that a human being with a slide rule cannot juggle with too
ny figures, a simple rule-of-thumb method of attempting to use 50%
overed metal in each melt was evolved. This method worked toler-
y well. Over a period investigated it showed that metal in category
was not used and that the amount of recovered metal available (in
A, B, and C categories) as a proportion of new melts required was
2%, and the actual amount of recovered metal used was 42·4%,
92% of that available. Linear programming was then applied to
if it could do better.

e Development of the Mathematical Model

f we define economic efficiency in this case as prescribing the make-
of a required alloy in such a way that the cost is at a minimum,
obtain a criterion for minimisation, provided that we can attach a
t to the use of the various recovered and virgin metals. Now the
t of one pound of virgin metal is its market price. Any recovered
tal in store has in fact two alternative uses: (*a*) it can be used in the
duction of alloys, or (*b*) it can be sold to outside firms such as metal
ners. The price a given recovered metal fetches, when sold outside,
ends on its tin content and its lead contamination, and roughly

speaking is 93–96% of the market price of tin. Therefore, the func that we have to minimise can be determined.

For a given requirement of a specified alloy on any given day can choose among the recovered metals available that day, with restriction that we cannot, of course, use more of any metal than fact available. Hence, taking the total amount of each recovered m available (in lb.) as unity, this restriction yields a set of n inequali all of which are of the form $x_i \leqslant 1$, where n is the number of recove metals available and x_i is the amount of the i'th recovered meta be used.

The amounts of tin, copper, antimony, and lead in each recove metal used, plus the virgin metal needed to add up to the amo required in the specification, furnish three equations and one m inequality.

Thus for tin:

$$t_1x_1 + t_2x_2 + t_3x_3 + \ldots + t_nx_n + x_{(n+1)} = T$$

where t_1 is the amount of tin in the x_1 recovered metal used,

t_2 is the amount of tin in the x_2 recovered metal used,

t_3 is the amount of tin in the x_3 recovered metal used,

t_n is the amount of tin in the x_n recovered metal used,

$x_{(n+1)}$ is the amount of virgin tin added,

T is the total amount of tin required in the alloy.

Similarly for antimony

$$a_1x_1 + a_2x_2 + a_3x_3 + \ldots + a_nx_n + x_{(n+2)} = A$$

for copper

$$c_1x_1 + c_2x_2 + c_3x_3 + \ldots + c_nx_n + x_{(n+3)} = C$$

and for lead

$$l_1x_1 + l_2x_2 + l_3x_3 + \ldots + l_nx_n + 0{\cdot}000\,7x_{(n+1)} \leqslant L$$

The coefficient 0·000 7 of $x_{(n+1)}$ is due to the fact that commer virgin tin contains 0·07% lead.

The function to be minimised is

$$Z = \Sigma(m_ix_i)$$

where m_i is the cost of the i'th recovered or virgin metal and x_i amount of recovered or virgin metal used in the prescription ($0 \leqslant x_i$ except when i refers to virgin metal).

Thus there are three equations, $n + 1$ inequalities, and $n + 3$ knowns (the amounts of each recovered metal to be used and quantity of each virgin metal required).

To convert the inequalities into equations in order to apply 'Simpl methods one new variable must be added for each inequality. If th

riables are denoted by x_i' the complete set of equations is then as en below:

$$x_1 + x_1' = 1$$
$$x_2 + x_2' = 1$$
$$\ldots\ldots\ldots\ldots$$
$$x_n + x_n' = 1$$
$$t_1x_1 + t_2x_2 + t_3x_3 + \ldots + t_nx_n + x_{(n+1)} = T$$
$$a_1x_1 + a_2x_2 + a_3x_3 + \ldots + a_nx_n + x_{(n+2)} = A$$
$$c_1x_1 + c_2x_2 + c_3x_3 + \ldots + c_nx_n + x_{(n+3)} = C$$
$$l_1x_1 + l_2x_2 + l_3x_3 + \ldots + l_nx_n + 0{\cdot}000\,7x_{(n+1)} + x'_{(n+4)} = L$$

In general mathematical usage it is not necessary to assign a meaning the variables added to convert inequalities into equations. In linear ogramming it is always possible to assign a meaning, and in this mple x_i' means *not* to use x_i' of the i'th recovered metal and $x'_{(n+4)}$ ans *not* to use the full amount of lead allowable. The set of equations cribing the restrictions of the problems thus consists finally of $n + 4$ iations in $2n + 4$ unknowns.

There is another way to arrive at this set of equations which makes asier to obtain the model: namely forming a so-called activity table. e use of any recovered metal can be considered as a process which has input, i.e. the recovered metal; and an output, i.e. the various ele- nts which it contains. Using a metal is an 'activity'; the various uts and outputs are 'commodities'. The activity table is formed by ing all possible activities. It is a convention to enter inputs with gative signs, outputs with positive signs. The nomenclature is a con- tion derived from economic theory (see reference 2). The activity le which forms our model is shown in Table I.

f to this table we add a last column, representing the restrictions of problem, which gives the availabilities of the input commodities l the requirements of the output commodities, the vector so obtained tains all the restrictions of the problem. It is called the boundary equirements vector, and its usual notation is P_0. If we now imagine uality and/or inequality signs between this vector and the last umn of the activity table, and plus signs between each column in h row of the table, we derive the set of equations and inequalities ch describe the problem. Again each activity has a known cost m_i, l we wish to find a solution to our set of equations and inequalities erms of x_i, $i = 1, 2 \ldots (n + 3)$, which minimises the function Z. t now remains to show that a 'feasible', i.e. non-negative, solution ts. The first step is to convert the inequalities into equations by addition of other variables, i.e. by adding activities of a certain n. These new activities, known as disposals, have in our case the aning *not* to use the recovered metals to which they refer, and the one (column $P'_{(n+1)}$ of Table II) means *not* to fulfil the lead uirements, i.e. to have less than the maximum lead contamination

TABLE I

Notation ..	P_1	P_2	P_3	...	...	...	...	P_n	$P_{(n+1)}$	$P_{(n+2)}$	$P_{(n+3)}$		P_0
Activity / Commodity	Use x_1	Use x_2	Use x_3	...	...	...	...	Use x_n	Use Virgin Tin $x_{(n+1)}$	Use Virgin Antimony $x_{(n+2)}$	Use Virgin Copper $x_{(n+3)}$		Boundary
x_1	-1											$\geqslant$	-1
x_2		-1										$\geqslant$	-1
x_3			-1									$\geqslant$	-1
⋮				⋱									⋮
⋮					⋱								⋮
⋮						⋱							⋮
⋮							⋱						⋮
x_n								-1				$\geqslant$	-1
Tin	t_1	t_2	t_3	...	...	...	...	t_n	1			$=$	T
Antimony ..	a_1	a_2	a_3	...	...	...	...	a_n		1		$=$	A
Copper ..	c_1	c_2	c_3	...	...	...	...	c_n			1	$=$	C
Lead	l_1	l_2	l_3	...	...	...	...	l_n	0·000 7			$\leqslant$	L

Notation

x_i = Amount of i'th recovered melt ($i = 1, 2, \ldots, n$)
t_i = Tin content of i'th recovered melt
T = Tin content of required alloy
A = Antimony content of required alloy

TABLE II

Notation ..	P_1	P_2	P_3	...	...	P_n	$P_{(n+1)}$	$P_{(n+2)}$	$P_{(n+3)}$	P_1'	P_2'	P_3'	...	...	P_n'	$P'_{(n+1)}$	P_0
Activity / Commodity	Use x_1	Use x_2	Use x_3	...	...	Use x_n	Use Virgin Tin $x_{(n+1)}$	Use Virgin Antimon $x_{(n+2)}$	Use Virgin Copper $x_{(n+3)}$	DO NOT use x_1'	DO NOT use x_2'	DO NOT use x_3'	...	...	DO NOT use x_n'	Unused Lead Allowance $x'_{(n+4)}$	Boundary
x_1	−1									−1							−1
x_2		−1									−1						−1
x_3			−1									−1					−1
⋮				⋱									⋱				⋮
⋮					⋱									⋱			⋮
x_n						−1									−1		−1
Tin ..	t_1	t_2	t_3	...	...	t_n	1										T
Antimony	a_1	a_2	a_3	...	...	a_n		1									A
Copper ..	c_1	c_2	c_3	...	...	c_n			1								C
Lead ..	l_1	l_2	l_3	...	...	l_n	0·000 7									1	L

TABLE III

Notation / Commodity	$P_{(n+1)}$	$P_{(n+2)}$	$P_{(n+3)}$	$P'_{(n+1)}$	P_1'	P_2'	P_3'	...	...	...	...	P_n'
x_1					-1							
x_2						-1						
x_3							-1					
⋮								⋱				
⋮									⋱			
⋮										⋱		
⋮											⋱	
x_n												-1
Tin	1											
Antimony		1										
Copper			1									
Lead	0·000 7			1								
Feasible solution ..	T	A	C	L $-0{\cdot}000\,7T$	1	1	1	1	1	1	1	1

wed. These disposals have zero cost. The system of equations ined is set out in Table II.

Vhen in linear programming a set of inequalities is converted into ations by the addition of such disposal vectors, it is always possible ssign a meaning to them, though they are not necessarily always ero cost.

he next step is to find a feasible solution. In this case it is clear by using virgin metals only for a given required alloy a non- ative solution to the set of equations in Table II is obtained. The rix of coefficients of this solution is given in Table III.

his solution in explicit form is

$$x_i = 0 \quad i = 1 \text{ to } n$$
$$x_i' = 1 \quad i = 1 \text{ to } n$$
$$x_{(n+1)} = T$$
$$x_{(n+2)} = A$$
$$x_{(n+3)} = C$$
$$x'_{(n+4)} = L - 0{\cdot}000\,7T$$

s matrix must now be inverted, clearly a very simple operation in case. From this stage onwards 'Simplex' computation is used.

Ve have now shown that our problem is amenable to treatment by ar programming methods, since a feasible solution can be found.

Trial Run

he efficacy of the method was tested by a trial run. Prescriptions he past eighteen months were available, which showed that under old methods recovered metals in category D tended to accumulate. consecutive weeks were selected at random, and recalculated by linear programming method. Comparison of the results showed with linear programming D-category recovered metals again led to accumulate, e.g. the same type of recovered metal was cted for re-alloying by both methods, and recovered metals of cate- D are not suitable for re-use.

he saving due to linear programming was comparatively small and erms of recovered metal used is shown below. The period investi- d was 10 weeks and the amounts are expressed as a percentage of and.

Usable recovered metal (A, B, C)	46·2%
Recovered metal used by old method	42·4%
Recovered metal used by linear programming . .	44·6%

should here be pointed out that linear programming did not give criptions at the absolute minimum cost (46·2%) because of the ctical restriction that the foundry can weigh out metal only to the est pound, and that it is impracticable to split a recovered metal too small quantities.

The percentage of recovered metal used in a prescription by li programming showed wider fluctuations than the percentage u the old method (Fig. 1). This was to be expected, because the

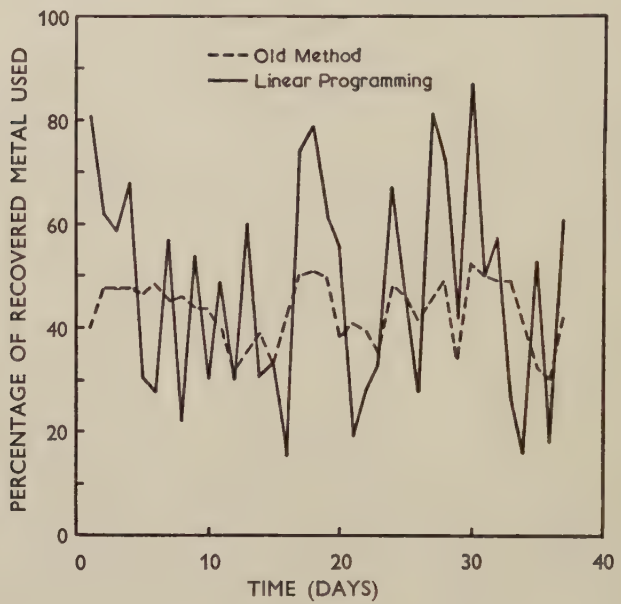


Fig. 1. Fluctuation in percentage of recovered metal used.

method tried to use a given percentage, 50%, of recovered met each prescription, irrespective of the amount of recovered metal a able at the time the prescription was made, whereas linear progr ming always uses the maximum amount available at any time.

Linear programming speeded up the turnover of recovered m in the store. This is an advantage when storage space costs mo Table IV and Fig. 2 show the effect of linear programmin accelerating the turnover.

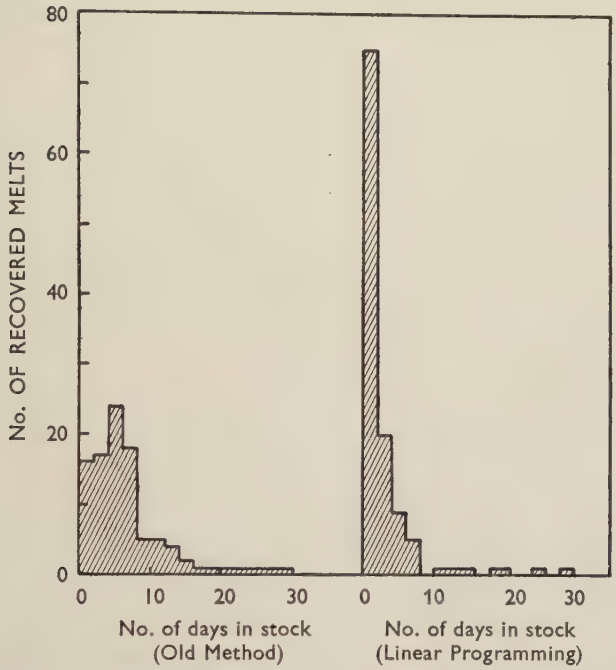
TABLE IV

'Average weekly stock' of recovered metals in thousands of pounds (the figures are coded and represent differences, not actual stocks)

Old Method			Linear Programming		
Category:			Category:		
A	B	C	A	B	C
113	23	82	48	3	44

Linear programming induces greater fluctuations in the demand virgin tin. This at first sight seems to be a disadvantage, but, in all, virgin tin is used. In actual working the saving in virgin tin is, u

, greater than could be expected from the trial run, and the in-
ıse in the fluctuation is only of the order of 15% of the standard
r.



2. Frequency distributions of the number of days that recovered metals spent in stock.

inally, whereas under the old method it was practically impossible
onsider all the recovered metals available when prescribing, linear
gramming assures that this is done.

here is a slight theoretical difficulty in applying linear program-
g. In the model we assumed that at any time any recovered metal
in the solution would be sold. In fact such recovered metals
ain in store and become available for later prescriptions. Consider-
however, that our aim is minimising cost of prescription, and that
act we defined cost as opportunity cost, as that definition is used
conomic theory, it will be seen that the non-sale of recovered metals
in the solution does not really upset our model. Furthermore, the
stigation showed that certain recovered metals (category D) could
old safely whenever required, since it was not profitable to use
n for re-alloying.

very often happens that prescriptions for two or more alloys are
anded on the same day. For example, 1 000 pounds of specification
10 antimony, 5 copper, 0·02 lead, balance tin) and 2 000 pounds
ecification B (8 antimony, 3 copper, 0·02 lead, balance tin). The
el can easily be extended to cover this case, as shown in Table V.
to the size of this type of matrix the computation becomes rather
onged, even for only two alloys. For this reason, and because

TABLE V

Commodity \ Notation		P_1	P_2	P_3	P_4	P_5	P_6	...	...	...	...	$P_{(2n-1)}$	P_{2n}
x_1		-1	-1										
x_2				-1	-1								
x_3						-1	-1						
⋮								⋱	⋱				
⋮										⋱	⋱		
x_n												-1	-1
Alloy A	Tin	t_1		t_2		t_3		...		...		t_n	
	Antimony ..	a_1		a_2		a_3		...		...		a_n	
	Copper	c_1		c_2		c_3		...		...		c_n	
	Lead	l_1		l_2		l_3		...		...		l_n	
Alloy B	Tin		t_1		t_2		t_3		...		...		t_n
	Antimony ..		a_1		a_2		a_3		...		...		a_n
	Copper		c_1		c_2		c_3		...		...		c_n
	Lead		l_1		l_2		l_3		...		...		l_n

TABLE V (*continued*)

Commodity \ Notation		$P_{(2n+1)}$	$P_{(2n+2)}$	$P_{(2n+3)}$	$P'_{(n+1)}$	$P_{(2n+5)}$	$P_{(2n+6)}$	$P_{(2n+7)}$	$P'_{(n+2)}$	P_1'	P_2'	P_3'	...	...	$P_{n'}$	P_0
x_1										−1						−1
x_2											−1					−1
x_3												−1				−1
⋮													⋱			⋮
⋮														⋱		⋮
x_n															−1	−1
Alloy A	Tin	1														T_1
	Antimony ..		1													A_1
	Copper			1												C_1
	Lead	0·000 7			1											L_1
Alloy B	Tin					1										T_2
	Antimony ..						1									A_2
	Copper							1								C_2
	Lead					0·000 7			1							L_2

Notation as in Table I, but
T_1, A_1, C_1, L_1 = Amount of tin, antimony, copper, lead, required in alloy A
T_2, A_2, C_2, L_2 = Amount of tin, antimony, copper, lead, required in alloy B

prescriptions which are demanded at about 10.30 a.m. have to issued by 4 p.m. to the foundry, the following short cut has b adopted: A new alloy, C, which is the sum of the matrix specificat of A and B, is stipulated, and the programme is set up as if only a C were demanded (Table VI). After the solution to this has b

TABLE VI

Specification for alloys; 1 000 *lb. of A and* 2 000 *lb. of B required*

	Alloy A (%)	Alloy B (%)	Alloy A (lb.)	Alloy B (lb.)	Notional Allo C = A + B (lb.)
Tin	84·8	88·8	$848(T_1)$	$1\,776(T_2)$	$2\,624(T_1 + T$
Antimony ..	10·0	8·0	$100(A_1)$	$160(A_2)$	$260(A_1 + A_2$
Copper ..	5·0	3·0	$50(C_1)$	$60(C_2)$	$110(C_1 + C_2$
Lead ..	0·2	0·2	$2(L_1)$	$4(L_2)$	$6(L_1 + L_2$

found, a programme of the type illustrated in Table V is set up, only those recovered metals which appear in the solution for alloy are used. Computing this second programme produces the requi solution A and B. It must, however, be said that this short cut n not always work, i.e. the second programme may turn out a solut to A and B which does not use recovered metals and virgin metal the same proportion as indicated by the solution to C, and there the cost of making A and B is higher than that of making C. Sho that happen, A and B would have to be re-computed from the ma of Table V, using all the recovered metals available. In practice to now the short cut method has always worked, but on theoret grounds there is certainly no reason why it should do so. In ac practice the work of splitting up C into the components of A an can be done by inspection and a new programme is not necessary.

When it was decided to use linear programming for prescrib another simplification was added. We know from the investiga that recovered metals of categories A, B, C only are usable and t they tend to be used within a very short time from their becom available. Also they are an inevitable by-product of the manufactur process. Therefore the cost of these recovered metals can be set at ze In the function ζ which we wish to minimise only those m_i appertain to virgin metals are different from zero. For the m_i referring to vir metals the computer can choose any numbers that are found c venient, provided only that their ratio to each other is the same as ratio of market prices of the corresponding virgin metals. With th simplifications the computing time has been cut down considerab On the average less than two hours per day are spent on comput prescriptions. The device of using computationally convenient nu bers, instead of the actual m_i, can always be used if the actual maxi

minimal) value of Z is not required. In our case the simplifica-
n really means that instead of minimising the cost function given
lier, the function minimised is the amount of virgin metal used.
A further simplification, suggested recently by S. Babik of the
acier Metal Company, is that with many alloys where the limiting
d content is reached before the antimony, copper, and tin require-
nts are satisfied, the function minimised is the cost of reducing
$_{(+4)}$ to zero.

ganisation

The results of applying the linear programme do not appear to be
ctacular. They represent a use of some 4% more recovered metal.
is, together with the more rapid turnover of recovered metal stocks,
useful advance, however, and well worth the effort. The investiga-
n has provided a reminder that a well-tried rule-of-thumb method
industry can approach the theoretical optimum very closely and it
y well be that linear programming tried on a larger scale will result
the same order of useful but marginal improvements.

The initiation of the linear programming presented some practical
ficulties which other users of the method may have experienced.
t all the new melt-prescriptions asked for were straightforward, and
users of the previous rule-of-thumb method had acquired a skill
d an instinct for dealing with unusual situations. It was found
irable to make the transition very gradually so that these skills
uld be transferred. Where communication with the factory is con-
ned personal contact is very important, and the transition period
ped to establish the new contacts.

The number of recovered melts used for each new prescription had
merly rarely exceeded two; with linear programming it was usually
r and could have been as high as seven. This gave rise to some
tial handling difficulties. By regulating the receipt of incoming
mands the maximum number is kept to four. This results in a slight
s of mathematical efficiency but resolves the handling problems.

nowledgements

We are very grateful for the assistance of Dr G. Morton of the
ndon School of Economics; and to Mr C. A. Marr and Miss L.
rnberg of the Glacier Metal Company for their computational
rk. Thanks are due to the Glacier Metal Company for permission
ublish this paper.

REFERENCES

HARNES, A., COOPER, W. W. and HENDERSON, A. (1953). *An Introduction to Linear Programming.* Wiley, New York; Chapman and Hall, London.

OOPMANS, T. C., Editor (1951). *Activity Analysis of Production and Allocation.* Wiley, New York.

SAMPLING FOR VARIABLES WITH A VERY SKEW DISTRIBUTION

R. TULSE
Research Techniques Unit, London School of Economics

This article propounds a sampling method that may be use in certain types of survey work, and describes a test of efficiency when applied to an artificial population with kno characteristics.

Introduction

It frequently occurs in sample-survey work that the variable to studied is of particular intensity in a small sub-group of the populatio For example, in samples of human populations the purchase of co sumer durable goods takes a zero value for a large proportion of t community and a value of unity for the remainder. Other variabl such as cigarette smoking, take zero or very small values for a la proportion and high values for a few. It may be that this sub-gro is one on which particular interest centres; for example, if we estimating the mean consumption of cigarettes, we should like collect information from more of the heavy smokers. Any stand. method of random sampling from a complete sampling frame will g proportionally small numerical representation to the sub-group. If special sampling frame can be constructed, any attempt to incre total sample size so as to give a sufficiently large sample of the s group may result in the survey becoming impossibly costly.

However, there are circumstances in which a small sub-group can given increased numerical representation in the sample without p hibitive increases in cost and without imparting bias to the estimat procedures to be used. Such circumstances arise when the b information identifying the individual as a member of the sub-gro may be reliably obtained from 'neighbouring' members of the gene population. ('Neighbouring' is written with quotation marks beca it does not imply physical neighbourhood in general, although in t sampling of human populations 'neighbourhood' is often the hou hold.)

The advantages of a procedure of this kind are twofold. First, su sidiary studies of the sub-group are given an adequate sample si and, second, the special status of the sub-group may enable us, ir spective of the first advantage, to improve the accuracy of estimation the whole sample.

Below we outline a method based on the collection of informati from neighbouring members of a group. The procedure to be discus

biased: it will sometimes lead to decreased precision of estimation e sense of an increase in sampling variance for fixed sample size; t will almost always lead to a reduction in cost for fixed sample which will generally more than offset any increase in variance. procedure will thus, in general, be more efficient than ordinary om sampling in terms of variance per unit of cost.

Method Proposed: Theoretical Discussion

ppose that the population of individuals can be uniquely allocated useholds (or, in general, 'neighbourhoods') and that every member particular household possesses the basic identifying information in ect of all members of that household. The sampling method pro- d is as follows:

) Select a sample of households at random (with equal or unequal probabilities, with or without replacement).

) In each selected household select *all* individuals who are sub-group members and (at random) *one* individual who is not a sub-group member. If the household contains none in any of these categories, none can, of course, be included in the sample.

) For each variable under consideration, (*b*) above enables us to construct an unbiased estimate of the household mean value, since we know the household numbers in each category.

) Given these unbiased estimates of household means and the method of selection used in (*a*) above, we may use the Yates-Durbin rule[2] to construct an unbiased estimate of the population mean for each variable.

remarkable feature of the method, inherited directly from the s-Durbin rule, is that although the sampling variance of the esti- e of the population mean is complex, the estimator of this variance xtremely simple. The method generates correlations within the ple between sub-group members and others and also between sub- p members themselves, while the sample size in the sub-group gory, the sampling fraction in the other category, and the total ple size are all random variables. In the simple case of sampling n equal or unequal probabilities) with replacement, the complexity ne sampling within households has no effect on the form of the nator of sampling variance, which is simply obtained by inserting stimated household means in the estimator of variance applicable e first stage of sampling.

nus, for sampling m out of M households with replacement with abilities p_i proportional to the number of electors appearing in the toral Register, an unbiased estimator of the population mean is

$$\bar{X} = \frac{1}{Nm} \sum_{i=1}^{m} \frac{N_i \bar{X}_i}{p_i} \qquad \ldots . (1)$$

where N_i is the number of individuals and $\bar{X}_i$ is the estimated mea
the i'th household, and $N = \sum_{i=1}^{M} N_i$ is the total number of individu
the population. If, as is generally done, we take it that for all prac
purposes

$$p_i = \frac{N_i}{N}$$

so that we are now sampling households with probabilities proporti
to their sizes, (1) becomes

$$\bar{X} = \frac{1}{m} \sum_{i=1}^{m} \bar{X}_i , \qquad \ldots$$

the simple mean of the household estimated means. The samp
variance of $\bar{X}$ is estimated by

$$\hat{V}(\bar{X}) = \frac{1}{m(m-1)} \sum_{i=1}^{m} (\bar{X}_i - \bar{X})^2 \qquad \ldots$$

and $\hat{V}(\bar{X})$ is unbiased for $V(\bar{X})$. If, on the other hand, we sar
households with equal probabilities and put $p_i = 1/M$ in (1), we ob

$$\bar{X} = \frac{M}{N} \cdot \frac{1}{m} \sum_{i=1}^{m} N_i \bar{X}_i = \overline{N_i \bar{X}_i} / \bar{N} \qquad \ldots$$

where $\bar{N}$ is the average number of individuals per household, and

$$\hat{V}(\bar{X}) = \frac{1}{\bar{N}^2 m(m-1)} \sum_{i=1}^{m} (N_i \bar{X}_i - \overline{N_i \bar{X}_i})^2 \qquad \ldots$$

Equation (3) is evidently simpler than (5) and $\bar{N}$ will often be unkn
precisely. This first procedure is therefore recommended. In prac
application in the UK it is, in fact, the simpler to operate.

The Efficiency of the Method

Like ordinary two-stage sampling, the method proposed incr
sampling variance—over that obtained by simple random samplir
individuals—by sampling in correlated household clusters, the cor
tions usually being positive. The second-stage stratification and
per cent sampling of sub-group members counteract this by redu
variance. This, together with the reduction in costs to be expe
from concentrating interviewing at fewer households, suggests that
proposed method's efficiency may be greater than that of sir
random sampling in terms of variance per unit cost. Little more
be said in general concerning the efficiency of the method, s
obviously everything depends on the precise character of the sub-gr
that is singled out for special attention. In general, this sub-group
be characterised by certain values for a variable believed to be cl
related to the variables being studied, and the efficiency of the me
will depend on how justified this belief is.

ıe special case which is of some interest is where the variable acterising the sub-group is itself the variable under study. (For ıple, in studying cigarette-smoking one might characterise the group as those members of the population smoking more than ty cigarettes daily.) In this case it is clear that the chance of oved efficiency is increased. The question is related to the problem ɔtimum stratification studied by Dalenius.[1]

Experiment

ıe method was tested by constructing a highly skew distribution of riate x with frequency function $10\,000\,e^{-x}$. The necessary fre-ıcies can be read off from tables of the exponential function and listribution function is simply $10\,000(1 - e^{-x})$. From this distribu- which was grouped in intervals of width 0·5 (from $x = 0$ to 9), a sample of 1 500 values was drawn at random, using the s of Kendall and Babington Smith.

ıese 1 500 values form our 'population'. They were listed in order ccurrence and then grouped into 'households' with the following titution:

Households of	*Numbers of 'households'*	*Numbers of individuals*
1	50	50
2	125	250
3	120	360
4	100	400
5	50	250
6	20	120
7	10	70
	475	1 500

constitution corresponds approximately to that of the population e UK by household in 1951.

ıis population was now sampled (by random numbers with re-ement) in sets of 20. To any individual there corresponds a sehold', which is thus chosen with probability proportional to its In each household there were chosen all members having $x \geqslant 1{\cdot}5$ one (at random) of those having $x < 1{\cdot}5$. An unbiased estimator e mean $\bar{x}_h$ for the household is then

$$\bar{x}_h = \frac{m_1 x_1 + \Sigma x_2}{m_1 + m_2} \qquad \ldots\ldots(6)$$

re there are m_1 members $< 1{\cdot}5$ and x_1 is the selected member in group; and m_2 members $\geqslant 1{\cdot}5$ and x_2 typifies the members in that p. By means of the above procedure 200 sets of 20 were drawn and an for each set constructed according to equation (1). We then d the mean of these 200 values and their variance.

The mean and variance of the original distribution are each k
to have the value 1. The mean of our 200 values was 1·024 wi
estimated variance of 0·017 976 5 and a standard error of 0·198.

In making a comparison with purely random sampling we must
account of sample numbers. In this experiment 7 135 indivi
were chosen, and the average number per set for the 200 sets
therefore 35·675.* If such a number were chosen at random
sampling variance of the mean would be 1/35·675 = 0·028 0
The variance under our modified procedure, as already noted,
0·017 976 5. Compared with random sampling our method ha
efficiency of about 150 per cent; or, putting it the other way ro
our method achieves the same precision as random sampling, w
using only two-thirds of the sample size.

In practice, of course, the increase in efficiency may be different
so far as the cost of obtaining information resides largely in trave
to the household, the gain would be greater; but if there are cor
tions between members of the same household it might be lowe
The method is also subject to the objections sometimes levelled ag
interviewing more than one member of a household; but against
it avoids abortive calls. On the whole it appears to us that the me
would, in many circumstances, be safe to use and would ach
substantial economy.

REFERENCES

[1] Dalenius, T. (1952). 'The problem of optimum stratification in a special t
design', *Skand. Aktuar Tidskr.*, 61.

[2] Durbin, J. (1953). 'Some results in sampling theory when the units are se
with unequal probabilities', *J. R. Statist. Soc.*, B, **15**, 262.

* The expected number was 37·26. In the population, a proportion P (=0·2331)
values $\geqslant 1{\cdot}5$, and the average sample number is

$$20\left\{\sum_{r=1}^{7} p_r(1-P^r) + P\sum_{r=1}^{7} rp_r\right\}$$

where p_r is the probability that a household is of size r. Here the first summation is
and the second is 0·884; substitution of these values in the above expression gives the
37·26 for the average sample number.

ANALYSIS AND PRESENTATION OF THE RESULTS OF FACTORIAL EXPERIMENTS

ANTHONY G. BAKER*
ICI Plastics Division, Welwyn Garden City

In this article Mr Baker describes a systematic method of setting down and analysing the results of factorial experiments. The method is particularly suitable for use with a desk calculating machine and is applicable to either qualitative or quantitative factors.

s paper describes a method for analysing the results of factorial eriments which gives the mean square relating to each degree of dom. When the independent factors are quantitative, the method be extended to include the fitting of equations. Inversion of the lts, that is reconstruction of the yields of the individual factor binations, can be applied to experiments where the factors are er quantitative or qualitative; this computation is also described. he general approach to the analysis of factorial experiments follows es[1] and developments of his method are given by Eisenklam,[2] ies,[3] and Bainbridge, Grant and Radok.[4] The method given here nilar to these but has been found to be particularly useful when calculations are being carried out on a desk calculating machine. es[1] first described the method of inverting an analysis, and his nod is here extended to the general case. The main reasons for loping and adopting this method are:

it can be applied to any factorial experiment;

the computations are more convenient when using a desk computing machine, since the blocks of computation are of a more convenient size;

a computer can be trained to use the method in a comparatively short time;

the computation can be broken down into convenient stages;

the inversion of the results can be obtained using the same type of computation.

tion and a Definition of Standard Order

, B, C, . . . will be used to denote factors, combinations of a, b, c, will be used to denote the results of the experiments. Suffixes will sed to denote the level of the factors, thus for an n-level factor the

low with Unilever Ltd.

suffixes will be 0, 1, . . ., $(n-1)$. Using this notation, the result
2×2 experiment can be written as:

$(b_0\ a_0)$
$(b_0\ a_1)$
$(b_1\ a_0)$
$(b_1\ a_1)$

It is seen that in this order the suffixes, read in the scale of 2,
0, 1, 2, and 3. Thus the standard order of a 2^n experiment is an arra
ment of the results in an order such that the suffixes, read in the s
of 2, give 0 to (2^n-1). Similarly for a 3^n experiment, the stand
order has the suffixes going from 0 to (3^n-1) in the scale of 3. In m
factorial experiments the standard order is not so easily defined.
an example consider a $2 \times 2 \times 3$ experiment, the results in
standard order are:

Scale → 6	3	1
c_0	b_0	a_0
c_0	b_0	a_1
c_0	b_0	a_2
c_0	b_1	a_0
c_0	b_1	a_1
c_0	b_1	a_2
c_1	b_0	a_0
c_1	b_0	a_1
c_1	b_0	a_2
c_1	b_1	a_0
c_1	b_1	a_1
c_1	b_1	a_2

The suffixes read in the scale given go from 0 to $[(2 \times 2 \times 3) -$
The scale is easily obtained, since for an $m \times n \times p \times q$ experi
the scale read from the right is 1, q, $p \times q$, $n \times p \times q$.

This definition of the standard order has been found useful in ch
ing that the results have been written down in the correct order. Y
introduced the arrangement of results in standard order and ga
definition of the order. It is useful to have two different metho
order to check that the results are arranged correctly before sta
the computation.

Analysis

The aim of the analysis is to estimate the magnitude of the ef
of the factors being investigated and show whether the effects
significant. The equations relating the results with the indepen
factors are also derived from the analysis.

The first step in the analysis is to arrange the results in the stand
order. At this stage it is very useful to plot the results in this o
since the plot often shows the conclusions that can be expected
the analysis. If the variance of the results varies, this can somet

een from the plot. If there are any freak results, they can often be d from the plot and can be queried before the analysis is carried out. ne method of analysis given in this paper consists of calculating an y of terms; by multiplying each column of the array by the results nged in the standard order, the effects of the various factors can neasured.

Experiment

ne array for a 2^3 experiment is as follows:

sults

$_0$ a_0)	+1	−1	−1	+1	−1	+1	+1	−1	(0)
$_0$ a_1)	+1	+1	−1	−1	−1	−1	+1	+1	(0)
$_1$ a_0)	+1	−1	+1	−1	−1	+1	−1	+1	(0)
$_1$ a_1)	+1	+1	+1	+1	−1	−1	−1	−1	(0)
$_0$ a_0)	+1	−1	−1	+1	+1	−1	−1	+1	(0)
$_0$ a_1)	+1	+1	−1	−1	+1	+1	−1	−1	(0)
$_1$ a_0)	+1	−1	+1	−1	+1	−1	+1	−1	(0)
$_1$ a_1)	+1	+1	+1	+1	+1	+1	+1	+1	(8)

e bracketed numbers in the last column will be discussed later.)

nus the first column of the array acting on the results gives the of the results, the second column gives a measure of the effect of ne next column of B, and so on through the remainder of the effects C, CA, CB and CBA. As in Yates,[1] with the change in notation, ee that the columns acting on the results can be given in the formal essions:

$$\begin{aligned} \text{Total} &= (c_1 + c_0)\quad (b_1 + b_0)\quad (a_1 + a_0) \\ A &= (c_1 + c_0)\quad (b_1 + b_0)\quad (a_1 - a_0) \\ B &= (c_1 + c_0)\quad (b_1 - b_0)\quad (a_1 + a_0) \\ BA &= (c_1 + c_0)\quad (b_1 - b_0)\quad (a_1 - a_0), \text{ etc.} \end{aligned}$$

ne array for the 2^3 experiment, which will be referred to as $\alpha(2^3)$, be calculated, since there is an algebra such that

$$\alpha(2^3) = \alpha(2) \times \alpha(2) \times \alpha(2)$$

algebra has the following rule for multiplication:

$$\begin{array}{cc} p & q \\ r & s \end{array} \times \begin{array}{cc} P & Q \\ R & S \end{array} = \begin{array}{cccc} pP & pQ & qP & qQ \\ pR & pS & qR & qS \\ rP & rQ & sP & sQ \\ rR & rS & sR & sS \end{array}$$

same rule for multiplication applies whatever the size of the arrays . It can thus be seen that $\alpha(2^3)$ can be obtained from $\alpha(2)$, where is the array for a one-factor, two-level experiment. It is easily n that

$$\alpha(2) = \begin{array}{cc} +1 & -1 \\ +1 & +1 \end{array}$$

A $2^2 \times 3$ *Experiment*

$\alpha(N)$ is obtained by adding a column of $+1$ terms before the ortho polynomials[5] given for an N-level factor.

Thus $\alpha(3) = \begin{matrix} +1 & -1 & +1 \\ +1 & 0 & -2 \\ +1 & +1 & +1 \end{matrix}$ $\alpha(4) = \begin{matrix} +1 & -3 & +1 & -1 \\ +1 & -1 & -1 & +3 \\ +1 & +1 & -1 & -3 \\ +1 & +3 & +1 & +1 \end{matrix}$

As before, we have $\alpha(2^2 \times 3) = \alpha(2) \times \alpha(2) \times \alpha(3)$.

The order of multiplication is important, since it can be seen that

$$\alpha(2) \times \alpha(3) \neq \alpha(3) \times \alpha(2)$$

i.e. the algebra is non-commutative. It is essential to analyse the re in the way corresponding to $\alpha(2^2 \times 3)$ if the results are written d in the order $2^2 \times 3$. Thus A must be the 3-level factor and B ar both 2-level factors.

$$\alpha(2 \times 3) = \alpha(2) \times \alpha(3) = \begin{matrix} +1 & -1 \\ +1 & +1 \end{matrix} \times \begin{matrix} +1 & -1 & +1 \\ +1 & 0 & -2 \\ +1 & +1 & +1 \end{matrix}$$

$$= \begin{matrix} +1 & -1 & +1 & -1 & +1 & -1 \\ +1 & 0 & -2 & -1 & 0 & +2 \\ +1 & +1 & +1 & -1 & -1 & -1 \\ +1 & -1 & +1 & +1 & -1 & +1 \\ +1 & 0 & -2 & +1 & 0 & -2 \\ +1 & +1 & +1 & +1 & +1 & +1 \end{matrix}$$

(The bracketed numbers in the last column will be discussed l

Premultiplying by $\alpha(2)$ again we arrive at the array given b with the results arranged in the standard order.

Results

$(c_0\ b_0\ a_0)$	+1	−1	+1	−1	+1	−1	−1	+1	−1	+1	−1	+1
$(c_0\ b_0\ a_1)$	+1	0	−2	−1	0	+2	−1	0	+2	+1	0	−2
$(c_0\ b_0\ a_2)$	+1	+1	+1	−1	−1	−1	−1	−1	−1	+1	+1	+1
$(c_0\ b_1\ a_0)$	+1	−1	+1	+1	−1	+1	−1	+1	−1	−1	+1	−1
$(c_0\ b_1\ a_1)$	+1	0	−2	+1	0	−2	−1	0	+2	−1	0	+2
$(c_0\ b_1\ a_2)$	+1	+1	+1	+1	+1	+1	−1	−1	−1	−1	−1	−1
$(c_1\ b_0\ a_0)$	+1	−1	+1	−1	+1	−1	+1	−1	+1	−1	+1	−1
$(c_1\ b_0\ a_1)$	+1	0	−2	−1	0	+2	+1	0	−2	−1	0	+2
$(c_1\ b_0\ a_2)$	+1	+1	+1	−1	−1	−1	+1	+1	+1	−1	−1	−1
$(c_1\ b_1\ a_0)$	+1	−1	+1	+1	−1	+1	+1	−1	+1	+1	−1	+1
$(c_1\ b_1\ a_1)$	+1	0	−2	+1	0	−2	+1	0	−2	+1	0	−2
$(c_1\ b_1\ a_2)$	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1

(The bracketed numbers in the last column will be discussed l

s before, we see that the columns acting on the results can be given
ıe formal expressions:

$$\begin{aligned}
\text{Total} &= (c_1 + c_0) \quad (b_1 + b_0) \quad (a_2 + a_1 + a_0) \\
A_L &= (c_1 + c_0) \quad (b_1 + b_0) \quad (a_2 \qquad - a_0) \\
A_Q &= (c_1 + c_0) \quad (b_1 + b_0) \quad (a_2 - 2a_1 + a_0) \\
B &= (c_1 + c_0) \quad (b_1 - b_0) \quad (a_2 + a_1 + a_0) \\
BA_L &= (c_1 + c_0) \quad (b_1 - b_0) \quad (a_2 \qquad -a_0), \text{ etc.}
\end{aligned}$$

ıere A_L is a measure of the linear effect of the factor A
A_Q is a measure of the quadratic effect of the factor A
B is a measure of the linear effect of the factor B
BA_L is a measure of the interaction between A_L and B

nple

ı large experiments the calculation of the arrays and the compu-
ɔn appear tedious. However, the analysis can be conveniently
ken down into stages. This is particularly useful when qualitative
ɔrs are involved, since then it is frequently desirable to have the
ysis for the various levels of the qualitative factors separately.
he example given here will be used to illustrate the following
ıts:

the breakdown of the analysis into stages;
the checks which can be applied;
the inversion of the results;
fitting equations to the data.

ı plastics the tensile strength of a polymer is often obtained by
ılding a flat sheet of the material and then cutting strips from it,
ch are tested on a tensometer. The temperature and pressure used
ıoulding the sheet affect the measured tensile strength. In the
ımple the temperatures used are 150° C., 170° C. and 190° C.
ded 0, 1, and 2 respectively); the pressures used are 1 ton and
ns per square inch (coded 0 and 1 respectively); and two types of
ımer are used (coded 0 and 1 respectively). Since the type of poly-
used is a qualitative factor it is desirable to analyse the results
each polymer separately before combining them. The tensile
ıgths have been rounded off to the nearest 10 and for convenience
ded by 10 in order to simplify the computation.
he model of the equation to be fitted is

$$Y = \text{constant} + (A)a + (A^2)a^2 + (B)b + (BA)ba + (BA^2)ba^2$$

re Y is the tensile strength, a is the temperature, b is the pressure
the bracketed capitals are the coefficients of the respective terms,
(BA^2) is the coefficient of the term ba^2.
he equation may differ for the two polymers, and thus a different
ation may be required for each polymer. The analysis has therefore
ı arranged so that the results for each polymer are obtained sepa-
y before being combined in a complete analysis.

The tensile strengths have been plotted in standard order in Fi and the plot shows the factors that are likely to be significant. Th can be seen that the level of tensile strength for the two polyme different, and that the form of the relationship between the te strength and the factors may be different. The standard deviatio the tensile strength is known to be 10 (120 degrees of freedom) therefore the standard deviation of the (tensile strength/10) will b thus giving unit variance.

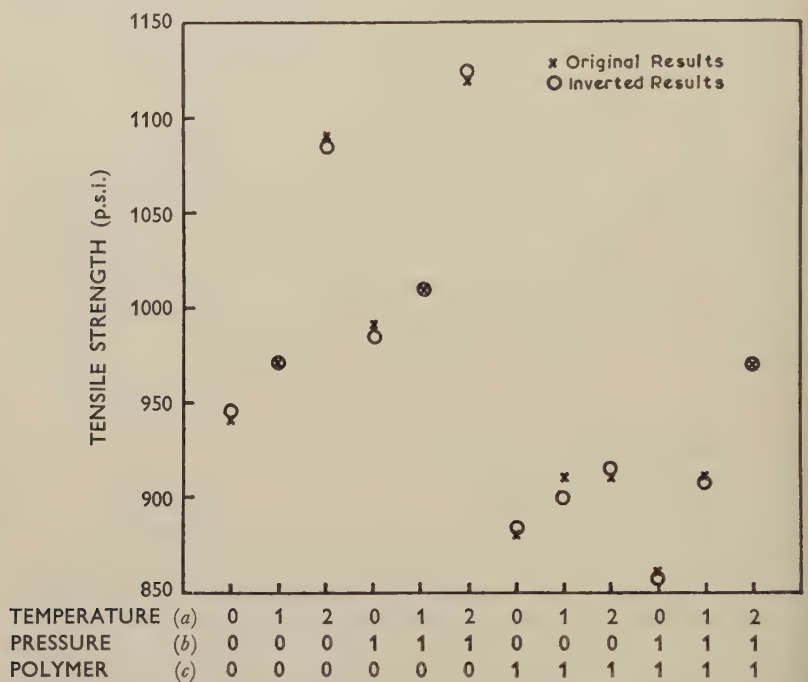

1150
1100
1050
1000
950
900
850
TENSILE STRENGTH (p.s.i.)
x Original Results
O Inverted Results
TEMPERATURE (a) 0 1 2 0 1 2 0 1 2 0 1 2
PRESSURE (b) 0 0 0 1 1 1 0 0 0 1 1 1
POLYMER (c) 0 0 0 0 0 0 1 1 1 1 1 1

FIG. 1. Tensile strength results plotted in standard order: experimental results and inv results.

The results are set down and analysed in the form shown by following Table, where *c* refers to polymer, *b* to pressure, and *a* temperature.

The first six results are operated on by $\alpha(2 \times 3)$, which is given the previous section. The six numbers obtained (612, 28, 18, 12, and 0) are written across the page, and these correspond to the to A_L, A_Q, B, BA_L, and BA_Q. For reference purposes the row will called the row E_0. Thus the first number in E_0 (612) is obtained operating on the first six results with the first column of $\alpha(2 \times 3)$, second number in E_0 is obtained by operating on the first six res with the second column of $\alpha(2 \times 3)$, etc. The calculation can checked by adding together the numbers in E_0, and the result 66 given in brackets. The value 668 can be checked as follows.

The sums of the terms in each row of the array $\alpha(2 \times 3)$ are cal lated; they are the bracketed numbers given with $\alpha(2 \times 3)$. It

Results									Checks	E	Σk^2	$E^2/\Sigma k^2$	Item
Standard Order	Tensile Strength ÷10												
$(c_0\ b_0\ a_0)$	94	E_0	612	28	18	12	−2	0	(668)	1 156	12	111 361·3	Total
$(c_0\ b_0\ a_1)$	97	E_1	544	14	−2	4	8	4	(572)	42	8	220·5	A_L
$(c_0\ b_0\ a_2)$	109	Σk^2	6	4	12	6	4	12		16	24	10·7	A_Q
$(c_0\ b_1\ a_0)$	99									16	12	21·3	B
$(c_0\ b_1\ a_1)$	101	$E_0^2/\Sigma k^2$	62 424	196	27	24	1	0	(62 672)	6	8	4·5	BA_L
$(c_0\ b_1\ a_2)$	112	$E_1^2/\Sigma k^2$	49 323	49	0·3	2·7	16	1·3	(49 392·3)	4	24	0·7	BA_Q
$(c_1\ b_0\ a_0)$	88		Total	A_L	A_Q	B	BA_L	BA_Q		−68	12	385·3	C
$(c_1\ b_0\ a_1)$	91	$E_0/\Sigma k^2$	102	7	1·5	2	0	0		−14	8	24·5	CA_L
$(c_1\ b_0\ a_2)$	91	$E_1/\Sigma k^2$	90·7	3·5	0	0·7	2	0		−20	24	16·7	CA_Q
$(c_1\ b_1\ a_0)$	86									−8	12	5·3	CB
$(c_1\ b_1\ a_1)$	91	Estimates	94·5	97·0	108·5	98·5	101·0	112·5		10	8	12·5	CBA_L
$(c_1\ b_1\ a_2)$	97		88·5	90·0	91·5	85·9	91·4	96·9		4	24	0·7	CBA_Q
												(112 064)	

thus be seen that (2 × 4th number) − (2 × 5th number) + (6 × number) equals the sum of the numbers in row E_0, i.e. 2(99) − 2(+ 6(112) = 668, which checks. The second six results are oper on by $\alpha(2 \times 3)$, and the numbers obtained are written in row these can be checked as before. In general the bracketed numb i.e. the sums of the rows, give a useful check. (The significanc rows E_0 and E_1 is considered in more detail later.) The colu of (E_0, E_1) are then operated on in turn by the first column of and then by the second column of $\alpha(2)$ to give the column E. T 1 156 = 612 + 544, 42 = 28 + 14, etc.; and −68 = −612 + 5 −14 = 14 − 28, etc.

Dr F. R. Himsworth has pointed out that the check column for $\alpha(2^2 \times 3)$ array can be obtained from the check columns for the α and $\alpha(3)$ arrays as follows:

Check column (C.C.) for $\alpha(2^2 \times 3)$

$$= \text{C.C. for } \alpha(2) \times \text{C.C. for } \alpha(2) \times \text{C.C. for } \alpha(3)$$
$$= (0 + 2)\ (0 + 2)\ (1 - 1 + 3)$$
$$= (0 + 2)\ (0 + 0 + 0 + 2 - 2 + 6)$$
$$= (0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 4 - 4 +$$

This can be generalised, and hence the check column for the array be obtained by the formal multiplication of the basic arrays. factors forming the product are written in the same order as in product of the arrays, since multiplication is again non-commuta

The E terms are then squared and divided by Σk^2, the sums of squares of the coefficients which multiply the results. It can be ca lated as follows.

For $\alpha(2)$ the sums of the squares of the terms in the column for total and linear effect are both 2. For $\alpha(3)$ the sums of the squar the terms in the columns for the total, linear, and quadratic effects 3, 2, and 6 respectively. This can be seen by writing down $\alpha(2)$ $\alpha(3)$.

$\alpha(2) =$	+1	−1	$\alpha(3) =$	+1	−1	+1
	+1	+1		+1	0	−2
				+1	+1	+1
$\Sigma k^2 =$	2	2	$\Sigma k^2 =$	3	2	6

The values of Σk^2 for the complete analysis are obtained by multiply the appropriate terms together.

Thus Σk^2 for $E(CBA_Q)$ = Σk^2 corresponding to the linear colu for $\alpha(2)$
× Σk^2 corresponding to the linear colu for $\alpha(2)$
× Σk^2 corresponding to the quadr column for $\alpha(3)$

$$= 2 \times 2 \times 6$$
$$= 24$$

re $E(CBA_Q)$ is the E term corresponding to the CBA_Q interaction. is result may also be obtained directly from the $\alpha(2^2 \times 3)$ array, y formal multiplication, cf. column checks.)

he E terms squared and divided by the appropriate value for Σk^2 given in the last column, which for convenience is called the $E^2/\Sigma k^2$ mn. The usual check can be applied to the numbers in this column, he sum of the numbers in the column equals the sum of the squares ie results on which the analysis is based. The sum of the squares of results is 112 064 and the sum of the numbers in the $E^2/\Sigma k^2$ column so 112 064. The terms in the $E^2/\Sigma k^2$ column are the mean squares, i having one degree of freedom, and it can be seen, by comparison i the residual variance of unity, that they are all significant at the level except BA_Q and CBA_Q. However, the important point is that e are large polymer interactions, and thus the results for each mer should be analysed separately. It is therefore necessary to ider the rows E_0 and E_1 in more detail. The appropriate numbers have been inserted, and the values of the mean squares are given he rows labelled $E_0^2/\Sigma k^2$ and $E_1^2/\Sigma k^2$. The sums of these mean ares can be checked as the column of mean squares was checked. Ve see that the mean squares for A_L, A_Q, and B are significant for first polymer and that the mean squares for A_L and BA_L are sig- ant for the second polymer. In order to examine these effects in e detail it is useful to obtain estimates of the original results and, quired, to fit the equations for the two polymers.

sion

o obtain estimates of the results, the values corresponding to $E_0/\Sigma k^2$ $E_1/\Sigma k^2$ are calculated for the mean square terms that are significant. erally if a term is not significant a zero is inserted, but if a term is significant and a higher-order interaction which contains the factor gnificant, then it is advisable to keep the term in, since this will give tter estimate of the original results. Thus for the second polymer have 0·7 for B, since A_LB is significant, even though B itself is not ificant.

he estimates of the results for the various combinations can then alculated. We will consider only the first polymer in order to illu- te the method. The numbers in the row $E_0/\Sigma k^2$ (replaced by zeros re necessary) are written in a column. These are then operated by the array:

$$\alpha^{-1}(2 \times 3) = \begin{matrix} +1 & +1 & +1 & +1 & +1 & +1 \\ -1 & 0 & +1 & -1 & 0 & +1 \\ +1 & -2 & +1 & +1 & -2 & +1 \\ -1 & -1 & -1 & +1 & +1 & +1 \\ +1 & 0 & -1 & -1 & 0 & +1 \\ -1 & +2 & -1 & +1 & -2 & +1 \end{matrix}$$

It can be seen that $\alpha^{-1}(2 \times 3)$ can be obtained from $\alpha(2 \times 3$ replacing the columns by the rows, e.g.

$$\alpha(2) = \begin{matrix} +1 & -1 \\ +1 & +1 \end{matrix} \qquad \alpha^{-1}(2) = \begin{matrix} +1 & +1 \\ -1 & +1 \end{matrix}$$

It is useful to plot the estimate obtained on the same graph as original results; this has been done in Fig. 1. If the equation is n good fit, either owing to the fact that a higher degree polynomia required or owing to the relationship not being expressible by a p nomial, this can sometimes be seen by comparing the originals with estimates. A contour map can be constructed from the estima though by using the equation this can be done more accurately.

Equations

Methods of fitting the equations to the data are given in referenc However, in my experience these can be confusing for a computer. T following rules give a straightforward method for fitting the equatio

(*a*) Orthogonal polynomials contain either all even-powered or odd-powered terms. Thus a quadratic orthogonal polynomial be of the form $A\eta^2 + B$.

(*b*) The coefficients of the polynomials are calculated as given bel Consider the quadratic term A_Q for the first polymer.

The values in the array $\alpha(3)$ for the linear column of A $-1, 0, +1$, and these are substituted for η. The expressions obtai are equated to the corresponding terms in the quadratic col of the array. Thus we have:

$$A(-1)^2 + B = +1$$
$$A(0)^2 + B = -2$$
$$A(1)^2 + B = +1$$

Only the first two equations are needed, and they give $A = 3, B = -2$, i.e. the orthogonal polynomial for A_Q is $3\eta^2$ -

(*c*) If we are considering a factor which has an even number of lev then

$$\eta = (x - \text{mid. value})/\tfrac{1}{2}(\text{interval length})$$

where x is the measure of the level of the factor.

If we are considering a factor which has an odd number of lev then

$$\eta = (x - \text{mid. value})/(\text{interval length})$$

For A_Q we have that the temperature levels were 150° C., 170° and 190° C. Thus if a denotes the level of temperature:

$$\eta = (a - 170)/20$$

Therefore the quadratic orthogonal polynomial for the temp ture is $3[(a - 170)^2/(20)^2] - 2$.

The polynomial obtained is then multiplied by the appropriate term of $E_0/\Sigma k^2$.

Thus for A_Q we have:

$$1{\cdot}5\{3[(a - 170)^2/(20)^2] - 2\}$$

Therefore if we write y = tensile strength/10, a = temperature, b = pressure, we obtain the equations:

First polymer

$$y = 102 + 7(a - 170)/20 + 4{\cdot}5[(a - 170)^2/(20)^2] - 3 + 2(b - 1\tfrac{1}{2})/\tfrac{1}{2}$$

Second polymer

$$y = 90{\cdot}7 + 3{\cdot}5(a - 170)/20 + 0{\cdot}7(b - 1\tfrac{1}{2})/\tfrac{1}{2} + 2(a - 170)(b - 1\tfrac{1}{2})/(20)\tfrac{1}{2}$$

If we write $Y = 10y$ = tensile strength, we have:

First polymer

$$Y = 990 + 3{\cdot}5(a - 170) + 0{\cdot}112\,5(a - 170)^2 + 40(b - 1\tfrac{1}{2})$$

Second polymer

$$Y = 907 + 1{\cdot}75(a - 170) + 14{\cdot}0(b - 1\tfrac{1}{2}) + 2(a - 170)(b - 1\tfrac{1}{2})$$

If we substitute the values used in the experiment we see that the equation gives the same estimates as those obtained by inverting the analysis.

owledgements

am indebted to Mr P. H. Hull (ICI Ltd Plastics Division) for the nal idea of using arrays in the manner described, to Mr R. H. dward (ICI Ltd Fibres Division) for suggestions on inverting the ysis, and to Miss M. H. Loraine (ICI Ltd Plastics Division) for king the paper.

REFERENCES

ES, F. (1937). *The Design and Analysis of Factorial Experiments.* Imperial Bureau of Soil Science, Harpenden.

NKLAM, P. (1953). 'Statistical methods in engineering experimentation', *Research*, **6**, 195.

IES, O. L., Editor (1954). *Design and Analysis of Industrial Experiments.* Oliver and Boyd (for Imperial Chemical Industries Ltd.), Edinburgh and London.

BRIDGE, J. R., GRANT, A. M. and RADOK, U. (1956). 'Tabular analysis of factorial experiments and the use of punch cards', *J. Amer. Statist. Ass.*, **51**, 149.

RSON, E. S. and HARTLEY, H. O., Editors (1954). *Biometrika Tables for Statisticians, Vol. I.* Cambridge University Press.

A PROBLEM IN PRODUCTION CONTROL

RAYMOND CHITTLEBURGH* and ALEC LEE†

In this article the authors show how simple statistical met can help to solve the problem of drawing up a produc schedule.

One of the principal intermediate products in the rubber industr rubberised fabric. Various combinations of different rubbers fabrics are called for in later stages of manufacture. The contro production of these processed fabrics, which are valuable, is import

In the past, production of rubberised fabrics has been carried ou short runs, and it has been possible to arrange the production sched from day to day. Although the Production Controller responsible this task has had a list of predicted requirements (made some t before) to guide him, essentially he has been able to use up-to- minute information about the state of the labour force, the condi of the equipment, and current requirements in drawing up schedule.

A development has been the introduction of a more or less tinuous method for processing fabrics which formerly were handle batches. The economics of continuous processing make it desirab have long production runs of processed fabric; because of the g diversity of fabrics, this could sometimes entail making at one enough of any type of rubberised fabric to satisfy the requirement some time ahead. The predictions of requirements have become m more important than they were, because the Production Contr has little other information available to help him plan his schec He cannot foresee, when making out his schedule, what deficienci labour or equipment may appear later.

When the possibilities of introducing continuous processing studied it became clear that a new production control procedure w have to be devised to assist the Controller, and the management st three conditions which this procedure should fulfil:

(*a*) It must allow for the (unavoidable) inaccuracy of the we forecasts of requirements for processed fabrics made by the plan in the 'consumer' departments.

(*b*) The risk of not being able to meet the requirements of the 'c sumer' departments at once must be small; on average this situa should not occur, for a given processed fabric, during more than weeks per year.

* Dunlop Rubber Co. Ltd., Research Centre.

† Formerly of the Dunlop Rubber Co. Ltd., now with British European Airways.

) It must produce stable orders; once a week's production schedule as been made in the fabric processing department it should not be hanged except in great emergencies, such as the situations referred in (*b*).

o begin with, the degree of accuracy of the weekly forecasts of uirements had to be assessed. Records were available, in each con- er department, of the forecast weekly requirements for processed ics for about four months; and (elsewhere) of the amounts actually l. The two were analysed statistically, and it was found that for specific class of processed fabrics there were periods of several ks when the forecast requirements remained almost constant, and actual requirements varied around them. During such periods it found that the relationship between forecast and actual require- nts could be written:

$$\text{(actual requirement)} = \text{(correction factor)} \times \text{(forecast requirement)} \quad \ldots\ldots(1)$$

re the correction factor, λ, had the distribution shown in 1. Tests showed that could be regarded as ormal distribution with ean of 0·85 for all prac- l purposes. Correlation grams were drawn to whether there was any al correlation between the ues of λ for consecutive ks. It was found that could be assumed that, ng periods of almost con- t forecast requirements, secutive values of λ were orrelated. Fig. 2 shows mples of the serial corre- on diagrams for one class rocessed fabrics.

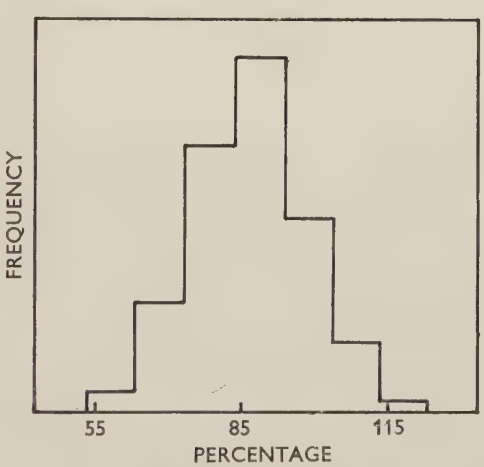


FIG. 1. Frequency distribution of processed fabric actually required expressed as a percentage of the corresponding forecast requirement.

rom time to time the level of the forecast—and actual—require- nts changed abruptly. These abrupt jumps were recognised as g due to economic factors whose effects were not predictable at time, but it was hoped that it would be possible to solve the prob- later. It was temporarily deferred because it was immediately ortant for several reasons, which cannot be given here, to devise oduction control scheme which would satisfy the three basic condi- s during periods when the forecast requirements remained constant. he knowledge of the relationship (1) and the form of the probability ibution of λ made it possible, given a weekly forecast of N rolls of

any processed fabric, to evaluate the probability $Q(k, N)$ that actual demand would exceed k rolls. The inaccuracy of the fore could then be allowed for in making out the production schedul the fabric processing department.

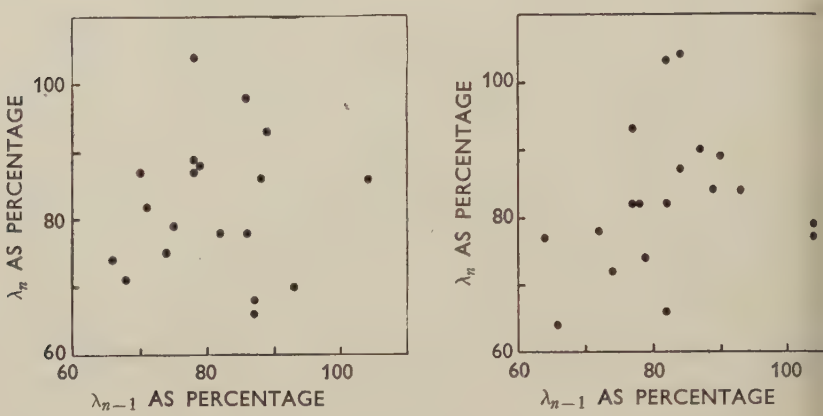


FIG. 2. Graphs of λ for the n'th week plotted against λ for the $(n-1)$'th week, for va values of n. (*a*) Processed fabric type I. (*b*) Processed fabric type II.

Secondly, it was clear from the nature of the distribution of correction factor that the requirement (*b*), in the absence of any st could be met only by making k rolls/week of each processed fa with k so large that $Q(k, N)$ would be less than 0·04 ($= 2/50$; t are 50 working weeks in each year). This would mean gross o production on average, with the creation of an uncontrolled st Consequently it was decided to create a controlled buffer stock of cessed fabrics.

Thirdly, if the buffer stock were controlled correctly, it would en the requirement (*c*) to be fulfilled. When actual demand was excee the scheduled output from the fabric processing department, deficiency could be made up from stock and the schedule need no altered until the end of the week. This would permit uninterru long runs of single processed fabrics. Thus it was clear that a sch could be based on the equation:

(Stock of i'th processed fabric [PF] at beginning of week $n + 1$)

= (stock of i'th PF at beginning of week n)

plus (amount of i'th PF produced during n'th week in reply to fore requirement)

minus (amount of i'th PF actually required during n'th week).

In symbols:

$$S_i(n+1) = S_i(n) + \alpha X_i(n) - \lambda_n X_i(n) \quad \ldots$$

re $S_i(n+1)$ and $S_i(n)$ are the stocks at the beginning of week $n+1$ week n, $X_i(n)$ is the forecast requirement for the n'th week, α is proportion of the forecast which should actually be made, and λ_n e proportion actually required.

ommon sense suggested that α must *at least* equal $\bar{\lambda}$, which is the n of λ and equal to 0·85; consideration of the quantity

$$S_i(n+1) - S_i(n)$$

ed that if α were actually equal to $\bar{\lambda}$, the stock level would fluctuate n alarming manner. During any given week there would be an l chance of a stock surplus or deficiency, and it is known that in cases it is possible for one alternative to occur far more frequently the other for long periods. In this particular situation, in order duce the risk of long-lasting cumulative stock deficiencies, either desirable stock level would have been much too great, or 'crash' rammes would have been required more often than specified by nanagement in condition (*b*).

is easier to cut back production than to increase it, and there was fore much to recommend the choice of $\alpha > \bar{\lambda}$. Equation (2) ed that with such a policy, since $\bar{\lambda} = 0{\cdot}85$, there would be on the age an overproduction of $(\alpha - 0{\cdot}85) \cdot X_i(n)$ units during week n, at under normal circumstances (small changes in X_i) the stock d tend to rise. However, by choosing α sufficiently large, the basic (mean stock level) needed to meet deficiencies could be kept l, and a re-ordering rule could easily be introduced to reduce excess l-up of stock. This will be described later.

ne amount of overproduction which would be desirable was rned by three factors: the size of basic stock which could be carried, ate at which a deficient stock should be restored to its correct size, the limit to the production rate of rubberised fabrics imposed by lant. In this particular situation the following conditions had to tisfied:

rst, the size of basic stock, for reasons of space and economy, could xceed thirty per cent of a normal week's output (or about twenty- per cent of a heavy week's). In general, if the demands for some of processed fabrics are high, those for the remainder are low, so the total basic stock of all types usually remains fairly steady, gh its make-up alters. Secondly, a high risk of actual requirements ding output of processed fabric for more than three consecutive s was undesirable. Finally, the processing rate was, and is, limited, ts potential is higher than that of subsequent production units.

was fairly easy to calculate the probability of actual demand ding the scheduled output (based on α times the prediction) for of n consecutive weeks. This is shown in Fig. 3, from which it was d that the amount of fabric scheduled for processing should be by $\alpha = 0{\cdot}89$, if the probabilities of the actual weekly requirement

exceeding the weekly output for two or three consecutive weeks
to be kept below the chosen limits. Other things being equal, it w
be permissible to take $\alpha = 0{\cdot}90$; that is, the amount of processed fa
produced in a week should be 90% of the estimated requirement

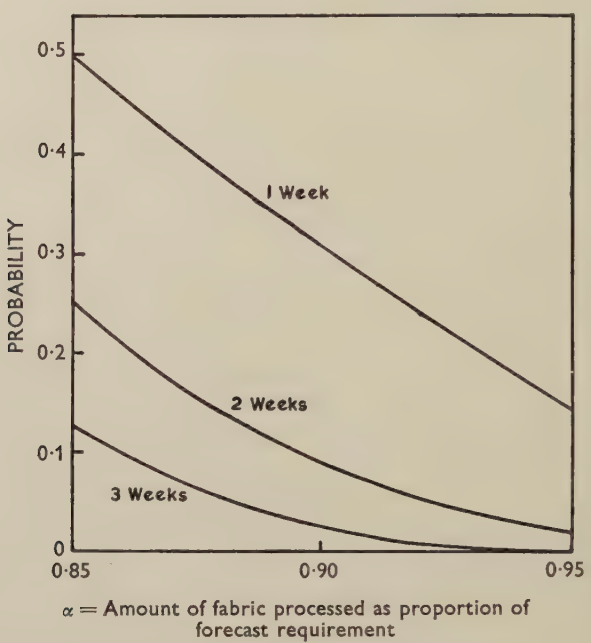


FIG. 3. Probability of the actual weekly requirement exceeding the weekly amount processed for one, two, and three consecutive weeks.

Now the probability P(n) that, over a period of n weeks, the
requirements would exceed the total produced during the period
the stock available at its commencement, is given by considering
distribution of the sum of n λ-variables, and referring to tables o
normal distribution. This is the risk that after n weeks insufficient
cessed fabric would be available to keep the consumer departm
working. If the basic stock level, S, were chosen so large that this
was small, then schedule adjustments would usually be unnecess
(In fact, they could always be made at the beginnings of weeks.)
was clear that the maximum risk would be only 0·013 with a s
level equal to 25% of the forecast requirement. This was quite acc
able to the management and the basic stock level was ch
accordingly.

It has already been mentioned that the system for supplying
cessed fabrics must also include some means of ensuring that the bu
stock should grow neither too large nor too small. This can be don
ensuring that there is an efficient feedback of information from
stockroom to the Production Controller.

ıe ideal stock level is 25% of the forecast requirement, and the ıuction Controller should aim to maintain this level. Consequently ı calculating his schedule for the next [the $(n+1)$'th] week he ld follow these steps:

) Calculate 90% of the forecast requirement.

$= \alpha X(n+1)$ units, where in this case $\alpha = 0{\cdot}90$]

) Check the stock and calculate the difference of the actual stock, $n+1)$, and the 'ideal'.

$= 0{\cdot}25X(n+1) - S(n+1)$ units]

If $[0{\cdot}25X(n+1) - S(n+1)] < 0$, there is excess stock, and the ıount of excess should be subtracted from the quantity calculated (*a*). Thus the amount of processed fabrics scheduled for produc-ɔn during the next [(n + 1)'th] week should be

$$90X(n+1) + 0{\cdot}25X(n+1) - S(n+1)$$
$$= 1{\cdot}15X(n+1) - S(n+1) \text{ units} \ldots . (3)$$

) If $[0{\cdot}25X(n+1) - S(n+1)] > 0$, there is a stock deficiency, ıd in order to restore the stock to its ideal level while at the same ne meeting the current actual demand, he should schedule

$$0{\cdot}90X(n+1) + K[0{\cdot}25X(n+1) - S(n+1)] \text{ units} \quad \ldots . (4)$$

will be clear that since there is, on average, overproduction when of the forecast requirement is produced, K need not be unity. The k level has been calculated so that it would tend to rise to its 'ideal' ion anyway. A value can be obtained for K by considering irement (*b*) and employing statistical distribution theory. In the ent case it was necessary to specify four values for K—one for each ur particular local conditions. It is clear that if normal working ısed on 90% of the forecast requirement, any increase to 92, 95 05%, in order to build up stock, must have repercussions on over-, bonuses, and so on. Thus K had to be given values which would v gradual build-up of stock; the penalty attached to an unneces-y rapid build-up would have been heavy. In effect, K is given the e zero unless the stock deficiency exceeds a critical figure, which lated to the production rate of the processed fabric concerned.

ıus a procedure was established for scheduling production against rs, and for controlling the stock. It was decided to test its effective-in two ways:

) By collecting the orders, and the records of the actual usage, for ch week during a period of months preceding the study and seeing ıether the new procedure would have proved more effective than e old.

) By collecting the current orders, week by week, calculating the obable requirements and the stock size, and arranging a schedule ıich could be compared with the actual operating results.

The first test was handicapped because the only past records avai were those which had been used in devising the new procedur appeared that during those weeks the new procedure would have tainly been an advantage over the old. The second test was ca out for about two months, and again the new procedure showe well. It was impossible for various reasons to prolong the second but from time to time checks were made of the λ-ratios for various t of processed fabrics and there was no indication of any change in distribution.

The problem of the sudden jumps in the level of requirements f particular type of processed fabric is caused by seasonal sales dem all over the world, for different products. It has not yet been sol Attention has, however, been directed to modifications of the prob described in this article.

One of us (A. M. Lee) has studied the effect of time-trends in values of λ, and has evolved a control procedure for use in situat where the distribution of λ is not stationary. It is hoped to desc this work at a later date.

Acknowledgements

We wish to express our thanks to the management of the Du Rubber Co. Ltd. for permission to publish this article.

USE OF REGRESSION ANALYSIS TO TECT ERRORS IN MEASUREMENT OF INTERMEDIATE MATERIALS IN A MULTI-STAGE PROCESS

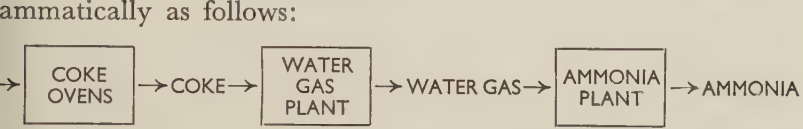
WILLIAM STANNAGE

Technical Department, Imperial Chemical Industries Limited, Billingham Division

Mr Stannage shows how certain unexpectedly wide variations in plant efficiency were found to be largely due to errors of measurement.

luction

the Billingham factory of ICI there are three important stages in nanufacture of ammonia from coal; they may be represented ammatically as follows:



the three stages the water gas plant is by far the most important at it is the most difficult to control. The raw material efficiencies e coke ovens and ammonia plant remain relatively constant (and es of variation are fairly apparent), but the efficiency of the water lant varies considerably, and since it is uneconomic to measure of the major losses of the process, sources of variation are not rent. A reliable index of the gas plant efficiency is of prime impor-, therefore, chiefly because any major raw material economies are to be made at that stage, but also because there is no easy way a carbon balance) of detecting and correcting errors.

Problem

e efficiencies of the three processes are defined as follows:

X (coke ovens efficiency) = tons coal/ton coke
Y (water gas plant efficiency) = tons coke/10^3 m.^{3} gas
Z (ammonia plant efficiency) = 10^3 m.^{3} gas/ton ammonia
Q (overall efficiency) = tons coal/ton ammonia

Hence $Q = XYZ$

will be noted that the so-called efficiencies are really reciprocal ncies.

the time of this investigation the recorded monthly efficiency of s plant had varied by $\pm$ 5 per cent from the mean in a period of ears; no suggested explanation of the month-to-month variation atisfactory, and hence errors in measurement were suspected.

The hypothesis was that, while the real values of Q, X, Y, and Z fairly constant, errors in measurement of coke and gas were ca substantial variations in the recorded monthly values of X, Y, an This seemed possible because coal and ammonia, as the raw ma and final product respectively, were both weighed; but the weig coke made (and consumed) was calculated from the weight and ana of the coal, making certain assumptions which were known t approximations (since the mechanism of carbonisation is not c pletely understood), while the quantity of gas made (and consur was estimated by an orifice-plate flowmeter of such large dimens that its characteristics had to be extrapolated from measurement considerably smaller ones. There was therefore at least an eleme doubt about the estimation of both intermediate products.

Detection of Errors in Coke and Gas Measurement

The hypothesis was borne out in the first place by the fact that coefficient of variation of the recorded values of Q (1·54 per cent) considerably less than the coefficient of variation of the recorded va of Y (2·58 per cent). If the true values of Q and Y were known, reverse must be the case, because the random variations of the ov efficiency would be made up of random variations in the three stituent efficiencies. The indication was therefore that the mea ments of coal and ammonia were considerably more accurate tha measurements of coke and gas.

Now let us suppose that over a period of months all true efficie remained the same, but that in one month there was an apprec error in gas measurement. The recorded values of Y and Z would change by the same fraction in opposite directions. By extension o argument, if frequent errors in the measurement of coke and occurred, a three-dimensional plot of Y against X and Z would hyperbolic relationships between Y and X and/or Z.

If we put

p = tons coal consumed/month
q = tons coke made/month
r = 10^3 m.^{3} gas made/month
s = tons ammonia made/month

Then

$$X = p/q$$
$$Y = q/r$$
$$Z = r/s$$
$$Q = p/s$$

If we take logarithms, the apparent relationships between log log Y, and log Z will become linear.

Putting

$$X' = \log X = \log p - \log q$$
$$Y' = \log Y = \log q - \log r$$
$$Z' = \log Z = \log r - \log s$$
$$(Q' = \log Q = \log p - \log s)$$

ıen if log q is in error by an amount δ, X' is in error by $-\delta$ and Y' $-\delta$, etc. Hence errors in measurement of coke or gas would lead to linear relationships between Y' on one hand, and X' and Z' on the r. Further, if the relationships arise solely because of errors in ;urement, a regression of Y' on X' and Z' should give slopes equal 1 in both cases, because equal and opposite effects are caused by one error. When variations arise in the true efficiencies, the rved efficiencies will show a scatter around a linear relationship. we write the regression equation as $Y' = a + bX' + cZ'$ the value is calculated to be $-0{\cdot}826$ with standard error 0·311, and the e of c is calculated to be $-1{\cdot}067$ with standard error 0·210.

ence both regression coefficients are significant and consistent with alue -1, and the hypothesis of errors in coke and gas measurement pported, provided that the true efficiencies are independent. A k on this is made later.

ıation of Magnitude of Errors

the correlations arise purely out of errors in measurement and the efficiencies are independent, then

$$\mathrm{E}[\mathrm{cov}\,(X'Y')] = -\mathrm{var}\,(\log q)$$

$$\text{and } \mathrm{E}[\mathrm{cov}\,(Y'Z')] = -\mathrm{var}\,(\log r)$$

e var $(\log q)$ and var $(\log r)$ are the *error* variances of $\log q$ and (the total variances of $\log q$ and $\log r$, if determined directly, would ıuch greater and would include variations in throughput and ency). Hence we have a way of estimating the magnitude of the s in q and r:

$$\mathrm{Var}\,(\log q) = [\mathrm{SD}\,(\log q)]^2$$

SD $(\log q)$ = coefficient of variation of q (expressed as a fraction)

om the covariances, the coefficients of variation of q and r were ıated to be 1·11 per cent and 1·76 per cent respectively; together would account for a coefficient of variation of Y of 2·08 per cent ıst the observed value of 2·58 per cent. Hence it is clear that the s in q and r contribute in a major way to the variations in Y, and two-thirds of the variance of Y is attributable to errors in q and r. ıermore, about three-fourths of the error variance arises from gas urement and one-fourth from coke measurement.

on Real Independence of X, Y, and Z

the value of either regression coefficient had differed from -1 by than twice its standard error, then there would have been a strong ımption that the dependence of Y on X or Z was not a result of thetical errors in measurement, but was a result of dependence on common variable. Even though the values of b and c were com- ly consistent with the error hypothesis, it was still consistent with lata to postulate that there were common variables which had

equal and opposite effects on the efficiencies of the water gas plant of the coke ovens and/or the ammonia plant.

In the case of the water gas and ammonia plant, for instance, com variables on which both efficiencies might depend were:

(i) gas throughput;

(ii) gas composition;

(iii) atmospheric temperature.

However, further regression analysis of X, Y, and Z on com variables disclosed no significant effects. Thus no satisfactory expl tion other than errors in measurement was found.

Conclusions

Where the product of one plant is the raw material of anot regression analysis of the logarithms of the efficiencies can be use detect errors in measurement of the common material, and to estir their magnitude, provided that the value of the regression coeffic can be shown to be consistent with the error hypothesis, and provi that the real independence of the efficiencies can be demonstrate

In the case described here, about two-thirds of the apparent va tions in recorded plant efficiency were shown to be the result of er in raw material and product measurement. Errors in gas (prod measurement contributed about three-fourths of the error varia and errors in coke (raw material) measurement about one-fourth

After the investigation the orifice-plate flowmeters (on which the measurement was based) were subjected to an intensive mainten effort, and it is satisfactory to record that during four years after investigation was originally made, no errors in measurement detectable by the same method.

Acknowledgement

The author wishes to thank Mr W. Spendley of the Statistical Se of ICI Ltd (Billingham Division) for the development of the ori idea into mathematical terms, and, indeed, for pointing out tha error variances of the measurements could be estimated from the variances of the logarithms of the efficiencies.

CALCULATING THE STANDARD ERROR OF A PROPORTION

EDWARD F. BRAYER
US Department of Labor

The author suggests the possibility of simplifying the arithmetical calculations needed to determine the standard error of a proportion and recommends an abridged approach where p does not exceed 0·1.

standard error of a proportion p is usually calculated by the well-
vn formula

$$\sigma_p = \sqrt{\left(\frac{pq}{\mathcal{N}}\right)}$$

e $q = 1 - p$ and $\mathcal{N}$ is the total population on which the estimate
s based. The formula may be rewritten $\sqrt{[(1 - p)n/\mathcal{N}^2]}$, where n
e number of observations whose proportion is being estimated.
n p is small, say less than 0·1, the formula may be approximated
$'(n)/\mathcal{N}$, with a proportional error of the order of $p/2$. Hence if p
or less the standard error is over-estimated by 5 per cent or less.

ample: 81 employees are injured out of a total of 1 032.

Hence p $= 81/1\,032 = 0{\cdot}078\,5$

Standard error $= \sqrt{(81)}/1\,032 = 0{\cdot}008\,7$

The correct value here is 0·008 4.

hen it is considered to be undesirable to approximate, the formula
be rewritten

$$\sigma_p = \sqrt{\left(\frac{n(\mathcal{N} - n)}{\mathcal{N}^3}\right)} = \frac{1}{\mathcal{N}}\sqrt{\left(\frac{n(\mathcal{N} - n)}{\mathcal{N}}\right)}$$

plication of the latter formula has been found to be advantageous;
rue that the number of steps in the calculation is not reduced, but
teps are easier and less likely to lead to error, because decimal
ions do not occur until the end. The use of a calculating machine
sumed. The example quoted above will be worked out by this
edure in order to demonstrate the layout of the calculations.

(1)	Record total number of observations	1 032
(2)	Show frequency of particular characteristic ..	81
(3)	Subtract (2) from (1)	951
(4)	Multiply (2) by (3)	77 031
(5)	Divide by (1)	74·64
(6)	Take square root of (5)	8·64
(7)	Divide (6) by (1)	0·008 4

e standard error of n is given by (6) and that of p by (7).

In a similar manner a method can be devised for the calculati the standard error of the difference between two proportions. usual formula is

$$\sigma_{p_1-p_2} = \sqrt{\left[pq\left(\frac{1}{N_1} + \frac{1}{N_2}\right)\right]}$$

where p and q are estimates obtained by pooling the estimates o two proportions, N_1 and N_2 being the total number of observatio each case.

The formula may be transformed into

$$\sigma_{p_1-p_2} = \sqrt{\left(\frac{(n_1 + n_2)\ (N_1 + N_2 - n_1 - n_2)}{(N_1 + N_2)\ N_1 N_2}\right)}$$

where n_1 and n_2 are the numbers in each estimate having a cer characteristic. This is suitable for easy computation with a machi

Example: One group has 1 829 observations, 202 having a cer characteristic; the second group has 1 789 observations, 240 ha this characteristic. Do the proportions in the two groups d significantly?

	Group 1	*Group* 2	*E* *Gr*
(1) Total number of observations in both groups	(1 829)	(1 789)	3
(2) Total frequency of characteristic	(202)	(240)	
(3) Subtract (2) from (1) ..			3
(4) Multiply (2) by (3)			1 403
(5) Divide by (1)			38
(6) Divide (5) by product of total numbers in each group, i.e. 1 829 × 1 789			0·000 11
(7) Take square root of (6) ..			0·0

The standard error of $p_2 - p_1$ is given by (7)

Now $p_1 = 202/1\,829 = 0{\cdot}110\,4$

$p_2 = 240/178\,9 = 0{\cdot}134\,2$

Therefore $\dfrac{p_2 - p_1}{\sigma_{p_2-p_1}} = 2{\cdot}18.$

The difference is significant at about the 0·03 level.

The method suggested here is believed to be much easier to tea particularly if those being instructed are not conversant with the u approach.

NOTES AND COMMENTS

:VELOPMENT which is of great interest to statisticians has been the forma- of the London Computer Group. The activities of the Group include neetings, addressed by leading personalities in the field of electronic al computers; (ii) the formation of sections to study various topics of est, such as administrative and financial considerations of installation, parative data on machines, training of personnel, etc.; (iii) the prepara- of information bulletins for members. Full details of the general meet- and study sections and application forms for membership may be ined from the Joint Honorary Secretaries, London Computer Group, Coleman Street, London, E.C.2.

★ ★ ★

mong the activities of the Union of Japanese Scientists and Engineers e publication (in English) of a journal, *Reports of Statistical Application rch*, which may be of interest to some of our readers. Although the to some extent invites comparison with *Applied Statistics*, perusal of two s gives the impression that the type of paper has closer similarities with in the *Journal of the Royal Statistical Society, Series B*.

★ ★ ★

achers of statistics and writers of books and articles often need to have rations of statistical naïvetés and fallacies. The following appeared in vincial evening paper:

"Fifty per cent of the population are below average intelligence", Dr G. Howells, a consultant psychiatrist, told the Royal Society of Health Ipswich, Suffolk, today. He said dullness and, less commonly, mental iciency, were largely responsible for ill-health and worry. Poverty, ld neglect, unpaid rents, and bad living conditions helped to produce problem family.'

★ ★ ★

e fourth Southern Regional Graduate Summer Session in Statistics e held from 12th June to 20th July 1957, at the Virginia Polytechnic ute, Blacksburg, Virginia. These sessions are of particular interest to search and professional workers who want intensive instruction in basic ical concepts and who wish to learn modern statistical methodology, achers of elementary statistics who want some formal training in rn statistics, (3) prospective candidates for post-graduate degrees in ics, (4) graduate students in other fields who desire supporting work tistics, and (5) professional statisticians who wish to keep informed of nced specialised theory and methods.

The staff for 1957 will include Daniel B. DeLury, J. L. McHugh, Evan J. Williams (visiting staff); and Willard O. Ash, L. S. Brenna, R A. Bradley, C. W. Clunies-Ross, John E. Freund, Rudolf J. Freund, Harshbarger, Clyde Y. Kramer, and R. Lowell Wine, from the Vir Polytechnic Institute.

The total tuition fee will be $38 for the six-week term. Doctoral cou will be offered to those holding doctoral degrees. Living and other expe at the Institute are reasonable. Inquiries should be addressed to H Harshbarger, Head, Department of Statistics, Virginia Polytechnic Insti Blacksburg, Virginia.

★ ★ ★

A new Industrial Applications Group has been started in Edinbu generally holding its meetings in the Heriot Watt College. It was addre in October by Mr K. T. Boyd on 'Statistics in Industry'; in Novembe Mr I. F. Hendry on 'Statistical Methods in the Paper Trade'; an December by Mr D. MacMahon on 'The Validation of Personnel Selec Tests'. Further meetings have been arranged, and the group, though s as yet, is enthusiastic. With Edinburgh as the background, its activities bound to be coloured by research laboratories and nationalised indus rather than by manufacture. We wish this new venture every success.

MEETINGS OF SECTIONS OF THE ROYAL STATISTICAL SOCIETY

INDUSTRIAL APPLICATIONS SECTION

BIRMINGHAM AND DISTRICT GROUP

October 1956. The Theory of Dams: D. G. Kendall

November 1956. Practical Difficulties in the Interpretation of the nalysis of Variance: Miss E. A. G. Knowles

December 1956. Practical Examples of Linear Programming: F. Benson

GLASGOW GROUP

November 1956. Time Study and Statistics: A. W. Swan

December 1956. The Statistics of Pass or Fail Tests: H. P. Stout

LEICESTER GROUP

October 1956. Market Research: D. H. Pickard

November 1956. Stock Control: J. H. Hepburn

December 1956. Statistical Calculations: A. Stafford Beer

LONDON GROUP

October 1956. The Detection of Periodicities in Time Series: An Appli- ion to Factory Sampling: R. C. Curnow

November 1956. The American Approach to Quality Control: C. S. ith

December 1956. Productivity Measurement at Plant Level: T. E. sterfield

MERSEYSIDE GROUP

October 1956. Simplified Statistics: J. S. Ogilvie

November 1956. Statistical Thinking in Industrial Control: A. Stafford r

December 1956. Planning Information from Linear Programming: C. Mackenzie

NORTH-EASTERN GROUP

October 1956. The Logic of Linear Programming: H. C. Mackenzie

November 1956. Statistical Treatments in Advertising Research: N. J. irrell

December 1956. Quality Control of a Continuous Process with Special ference to Paper Manufacture: I. F. Hendry

SHEFFIELD GROUP

25th October 1956. A Simplified Method of Regression Analysis: Miss W Gibson and G. H. Jowett

29th November 1956. Linear Programming in Industry: M. A. Aczel

SOUTH WALES GROUP

8th October 1956. Smoking and its Possible Effects on Lung Can W. R. S. Doll

This lecture was a report on the well-known work of Professor Bradf Hill and Dr Doll. The retrospective study, which gave sound informat as to the nature of the relationship between smoking and lung cancer, first described. Dr Doll then described the prospective study of member the medical profession, to which forty thousand doctors responded in 19 By 1956 the conclusions reached in the retrospective study were amply c firmed and the stability of the relationship is now well founded. A report has been published in the *British Medical Journal.*

12th November 1956. Some Statistical Problems arising from the Use Microspores in identifying Coal Seams: R. C. Tomlinson

10th December 1956. The Theory of Games: D. V. Lindley

STUDY SECTION

LONDON GROUP

10th October 1956. The New Retail Price Index: R. G. D. Allen

14th November 1956. Symposium on the Philosophy of Experimentat A. Bradford Hill, M. J. R. Healy, and M. G. Kendall

BRISTOL GROUP

18th October 1956. The New Index of Retail Prices: H. S. Booker

Mr Booker described the household and other surveys by means of wh data on expenditure had been collected, and discussed the problem coverage, rent, and seasonal purchases. He spoke of the atypical natur the 'average', but showed that the weights are relatively unimportant, gave examples of the difficulties encountered in dealing with change quality, controls, taxation, subsidies, and free goods.

15th November 1956. How Statistics can help Investigations in the Fact L. H. C. Tippett

Mr Tippett dealt with two types of investigation—surveys and exp ments. He gave examples of measures for taking account of uncontro variations in designing an investigation and interpreting the results, a discussed the influence on design of the need to avoid interference w production.

MEDICAL SECTION

30th October 1956. An Association between Smoking and Respirat Tuberculosis: C. R. Lowe

27th November 1956. The Aetiology of Malignant Disease in Childr D. Hewitt

4th December 1956. Medical Statistics at the National Institutes of Hea USA: H. F. Dorn

BOOK REVIEWS

Finney: *Experimental Design and its Statistical Basis*. Chicago: University of Chicago ss; London: Cambridge University Press, 1955. Pp. xi + 169. 30s.

new book by Dr Finney is always of interest, and this one is of a some- unusual nature. It is addressed primarily to research workers, espe- biologists, and its object is to make clear why it is useful to have the peration of a statistician both in the planning and in the analysis of an riment. The author points out that a statistician is often aware of certain le precautions which must be taken to ensure that the results of an riment will be reliable, in the sense that the effect under investigation t confused with other and irrelevant effects. Furthermore, it is often ble to modify a design so that it will give considerably more information ne same effort and the same cost. Such improvements often come from cussion between the statistician and experimenter, in which a better rstanding of each other's aims and methods may lead to developments h neither would have thought of alone.

arious chapters in the book explain the principal types of experimental n, such as randomised blocks, Latin squares, incomplete blocks, lattices, rial designs, confounding, and partial replication: but this is mainly to what possibilities are open to the experimenter, and there are no led instructions for the arithmetical analysis. There is also a discussion quential experimentation and of biological assays.

Finney illustrates his argument with a number of interesting and some- amusing results. For example, when three varieties A, B, C of headache dy were tried (by Jellinek), the respective proportions of cures were 0 and 80 per cent, figures which do not suggest any appreciable dif- ce in effectiveness. However, Jellinek was wise enough to include a *bo*, of no pharmacological activity, which nevertheless cured more than the headaches tested. Evidently these were of mainly psychological n: and when the subjects who responded to the *placebo* were omitted, it me clear that in fact treatment A was noticeably better than B or C.

is book is good propaganda for statisticians, and can be warmly recom- led to Unconverted Biologists. It should, however, be added that, ugh the author has taken great pains in writing the book, it is not an one to read. Dr Finney has tried to give the reader the benefit of his years of practical experience. He has distilled his wisdom down to sentials; but it is not too easy to grasp when expressed in a rather ral form and compressed into a moderately sized book. This would, rtunately, seem to be inevitable, though some of the examples will help ake the points clearer. Both the ordinary research worker and the ary statistician will benefit from reading this book, and if it leads to greater future collaboration, as it is intended to do, its publication ave been very well justified.

e book contains lists of elementary books on statistical analysis and ard textbooks on experimental design, as well as a rather miscellaneous ction of other references bearing on the subject. These lists will be help- those who wish to pursue the subject further. CEDRIC A. B. SMITH

E. P. Neale: *Guide to New Zealand Official Statistics (Third Edition)*. Auckland Univ College Texts No. 2. Christchurch: Whitcombe and Tombs, 1955. Pp. 195. 16s.

This third edition of a university textbook (first published in 1938) at being more than an unofficial guide to the official statistics of Zealand; it also discusses the general principles and problems of comp tion and includes some informed criticism of the official procedures in Zealand. The author is most at home with demographic statistics, which very fully covered. He describes not only the statistical problems but the administrative details of data collection, which is unusual but hel in such a short work. He also has much to say of interest on the de graphic history of New Zealand. The realities behind the statistics are ne lost to sight.

The treatment of retail price indices, although lengthy, is much less sa factory. There is no discussion of recent developments of sampling techni in the collection of household budgets. For instance, nothing is said of importance of a good sampling frame or of ways of securing good respo This incompleteness may be due to the fact that a satisfactory set of budg has apparently never been collected in New Zealand; the weights of retail price index are now based on estimates of aggregate supply of e commodity group (the methods used for this are not at all fully discusse

There is a more interesting discussion of the problem of compiling export price index in a country whose exports consist of a limited num of major primary products. It is shown, for example, how a unit value in with current weights, as formerly used in New Zealand, can rise, altho the unit value of each commodity separately falls (a curiosity which occur in the New Zealand export price index in 1931). The present export p index has fixed weights (exports in 1952).

Other sections are notable for their omissions. The discussion of w gives no indication whether the official indices relate to wage rates, earnings, or both. The discussion of national accounting is purely gen and says almost nothing about the New Zealand statistics. There is reference at all to the balance of payments statistics (as distinct from for trade statistics). The weakness of some sections is partly offset by a fasci ing description of the meteorological statistics, including a discussion problems of collection and comparability of records which are very sim to those encountered in economic and social statistics.

The author has succeeded in the far from easy task of imparting inte and significance to what could be a tedious description of statistical ser The material in this work, so far as it goes, is just what is needed guidance and reference by the university student, the research worker, the foreign user of New Zealand statistics. But it is seriously incomplete many ways, and bears all the signs of having been brought up to date v hastily. A thorough reorganisation, introducing some systematic arran ment of the material, separating definitions and comment, and with m attention to the basic facts about the matters under discussion, could t it into a model textbook for the applied statistician. C. T. SAUND

Enrique Chacón: Curso de Estadística: Volume I, Estadística Descriptiva y Modelo M mático. Bilbao: Editorial el Mensajero del Corazón de Jesús, 1955. Pp. xvi + 494.

As the title of this book implies, it is a treatise, in Spanish, on mat matical statistics. Some idea of its scope may be conveyed by a free tran tion of the chapter headings:

oncepts and definitions; Probability, basic postulates; Statistical data, ping, representation; Principal statistics of a frequency distribution; erlying mathematical model; Distribution functions for discrete vari- s; Distribution functions for continuous variables; Distribution functions as gamma, beta, F; Case of two variables, regression lines; Bivariate stics; Mathematical model for two variables, distribution functions; nsion to n variables; Time series; Correlation between time series; nents and other measures of various statistics (e.g. variance of mean, icient of variation, etc.); Statistics that follow a normal or asymptotically nal distribution; Statistics with non-normal distributions; Distributions her statistics, simplified tests; Multivariate statistics, generalisations of , F, etc.; Summary of formulae; Statistical tables and charts.

ttle remains to be said; the book appears to cover its chosen field ade- ely, and generally the presentation is good. From the practical stand- t it may seem a little odd that the analysis of variance makes only a appearance towards the end; perhaps detailed consideration may be ved for a further volume.

D. G. BEECH

e T. Landblom: *Nellie Landblom's Copybook for Beginners in Research Work*. Colorado ricultural and Mechanical College, Fort Collins, 1955. Pp. 118. $2.95.

ne title of this book is somewhat misleading, the sub-title 'Computa- al Procedures used in Applied Biometrics' giving a better indication of ontents, which consist largely of worked-out model examples covering y of the experimental design techniques in common use. It is necessary he reader to be familiar with the fundamental principles of statistics in r to know which model to use and how to interpret the results. The , which is reproduced from typescript, has been made available in nse to requests from graduate students.

hink that this type of book, which gives instruction in techniques with- he underlying principles, may be dangerous in the wrong hands. Hence not think that it deserves a 'public' wider than that which originally ested it.

D. G. BEECH

Dornbusch and C. F. Schmid: *A Primer of Social Statistics*. New York: McGraw-Hill, 5. Pp. vii + 251. 35s. 6d.

eviewers have recently tended to deplore the appearance of more and general textbooks in statistics instead of books dealing with special , and dealing with these more adequately. Certainly the special-field ook in statistical methods is needed, but perhaps there is also a need for s which aim to attract special groups to an interest in statistics generally, e hope that once attracted they may later go on to consider the more sticated methods that statistics can offer.

ne sociologist is a reluctant statistician, and this book states its target ence as social scientists and social workers. The preface states candidly the book makes no pretence to survey the entire field of statistics nor ve a thorough explanation of the topics covered. This self-confessed ation is fully borne out. The reason for producing the book is said to at it is unique in its presentation and orientation of the material. The of the book can therefore only be subjectively assessed. Is this original

presentation a fact? And if so, is it likely that it will have a greater ap to the social worker and the social scientist than other textbooks?

One way to gain the attention of reluctant students is to interest t from the outset. One way to interest them is to give them something they can realise is useful at the outset. To meet this point there are books that can compare with Snedecor's *Statistical Methods applied to Ex ments in Agriculture and Biology*. The present book does not touch on anyt which its readers are likely to regard as useful until nearly half-way thro its 250 pages. The approach is in the form we have become accustome expect in elementary statistical textbooks—tabular presentation, gra means, variances, ratios and rates; and thence to a treatment of distr tions, ending with correlation and contingency. The usual route to usual end. Quite satisfactory for some audiences, but there is little trac the 'original presentation' claimed.

Since social research workers—the target audience for this book— unusual among research workers mainly in that they are often in a posi to count but seldom to measure, perhaps this fact might have suggeste different presentation schedule. Even when social measurements are sible, very little can be said about the nature of the distribution. If then social research worker is to be interested in statistical methods he shoul interested in 'robust' statistics and non-parametric methods. Is there reason why the statistics of 'counting' should not be presented in the ea chapters? It should be possible, at least, after an introduction to the i of inference, to introduce χ^2 with several of its uses. Rank correlation c also be introduced before the end of the book. Fortified by these u statistics, which he may directly use, the social research worker might spurred to learn more techniques and to respect their limitations. He m also, perhaps, try to master a little mathematics.

This book fails to present to the social research worker any statis method which is not better presented in books which already exist. chapters on sampling, inference, and contingency are particularly poor.

LESLIE T. WIL

Seymour Melman: *Dynamic Factors in Industrial Productivity*. Oxford: Blackwell, 1956. xiii + 238. 22s. 6d.

Statisticians who have to do with industrial productivity are usually cerned with its measurement, but it is only right that they should int themselves in its improvement—the subject of the book under review. fessor Melman starts with two hypotheses: (1) that increased producti comes from increased mechanisation, and (2) that managements tend m and more to the adoption of mechanical methods of manufacture as relative costs of alternative methods, especially the cost of labour relat to that of machines, rise. It follows that high productivity is associated w high relative costs, and this relationship Professor Melman regards as basic dynamic factor in industrial productivity.

The second hypothesis is examined first by a costing investigation a inquiry in the British motor manufacturing industry. The costing investi tion showed that between 1938 and 1950 the cost of labour rose relative that of mechanical handling equipment, and the inquiry showed that had been accompanied by an increased use of such equipment, that factors other than the relative costs had changed, and that the hypoth

accordance with the way in which managements come to decisions. A llel examination of cost data for the US motor industry strengthens the for the hypothesis and shows also that in the ratio of labour to this icular machine-cost Great Britain is about ten years behind the US.

he first hypothesis is examined by correlating such national statistics as available in the literature of output per worker man-hour and the lled horsepower per worker (a measure of the degree of mechanisation). r the years, both have increased in the US; but, more significantly, n on the same scatter diagrams points are plotted for different countries, t countries each having one or two points for years in the 1930's and the having points for several years between 1899 and 1947, the correlation rprisingly high—at a glance it seems to be near to 0·9.

he combination of the two hypotheses is examined by studying the rela- ship between output per worker man-hour and the number of kilowatt- rs purchasable at the cost of a man-hour of industrial labour (a measure ie relative cost of factors, including machinery, other than labour). A er diagram of points for several countries at different times shows a elation coefficient of 0·88.

nother theme of the book is the increase in the ratio of administrative roduction personnel (horrid word!) employed in industry that has rred over the years in the UK and the US. This is not primarily due to increase in the number of designers and engineers accompanying in- sed mechanisation, and so Professor Melman regards it as diluting the ts of the increase in productivity of the production workers to such a ee that it may be a limiting factor to increased productivity.

his bald summary cannot do justice to the care with which Professor man develops and tests his argument, to the richness of his discussion hat happens in industry, and to the qualifications with which he states onclusions. The conclusions are broadly what anyone conversant with stry would expect, but it is valuable to have what would otherwise be ral impressions established after examination, and the process of exami- on arouses fruitful questions and thoughts. This book is a contribution ought and knowledge rather than a definitive treatment of its subject. influence of such factors as management techniques, management- er relationships, training, general attitudes of people to productivity, social conditions, is not considered. The conclusions of the book lead e view that these factors are of secondary importance—perhaps they be likened to lubricants and resistances to the motion of a machine r than to the energy that produces motion. But one would like to see examined with the thoroughness Professor Melman has brought to his tigation.

ave read somewhere that Swift failed in his object in writing *Gulliver's els* because what was written as a social satire has become only a ren's classic. I would not like to see a parallel fate for Professor nan's book, but nevertheless suggest that it has value for students of tics as an object lesson in the type of inquiry in which statistics plays t, and has to be combined with technical study of the field of applica- and cautious inference from mixed data. If the teacher disagrees with ssor Melman in some details of method or points of emphasis, that s the book all the more valuable from this point of view.

L. H. C. TIPPETT

J. M. S. Risk: *The Classification and Coding of Accounts.* Institute of Cost and Works Ac tants Occasional Paper No. 2. London: Gee and Co. (Publishers) Ltd. Pp. 79. 7s. (

Into these few pages go a set of principles for classification systen review of historical methods for the systematisation of knowledge, a ske code for the accounting classification, and some personal views on accounting as the true faith. The paper is a notable achievement in c pression, and most people would benefit from a study of the chapters dev to classification, symbolisation, and coding. In most sciences and pr sions classification systems have been built up by clear-thinking amate but the very diversity of business institutions has not helped the gen adoption of widely recognised terminology and classes of accounting d Dr Risk believes that the advent of the electronic store will force its use turn librarian, but while this is true so far as principle goes, there wil opportunity for yet more diversity in practice.

With the general chapters of the paper there can be little quarrel, tho one might have hoped for more discussion of the use of generally adop systems such as the Standard Industrial Classification or the United Nat Classification of Countries, and of the use of such systems even where s local disadvantage might occur. Many accountants, however, would themselves disagreeing with the author about his accounting system, this only illustrates his difficulty in attempting classification where thou is not crystal clear, and giving disproportionate space to contempo accounting controversies. It is surprising that no mention is made of national accounting system, where classification has perhaps gone fur than in other branches of the profession.

Although the author's application of his principles to his own wo admirable, it is sad to record that he offers no service on references. experiment on transcription-error determination reported on page noted with no acknowledgement; this is a pity, because the results w appear to need more justification than that given in the text.

C. NUT

Francis G. Cornell: *The Essentials of Educational Statistics.* New York: John Wiley and London: Chapman and Hall, 1956. Pp. xii + 375. 46s.

The 'Wiley Publications in Statistics' now comprise over thirty volu and the present book is a useful addition to the series. The author is fessor of Education in the University of Illinois, and the book shows hand of the experienced and clear-thinking teacher (though the ma matician would object to the setting out of the arithmetic test of form as given, for example, on page 50).

The first four chapters deal with the usual more elementary descrip statistics, up to skewness and kurtosis. After that the treatment is n modern, and sampling ideas, inference, 'power', and critical regions ar well dealt with. Although the book is primarily intended for education much of it sets out the general background in such a way that worker other disciplines would profit by it; in fact, such specifically psychom techniques as factor analysis are not included.

The later part of the book considers the sampling distribution of mean, with one- and two-tailed tests, and standard error; the resi variance in the case of regression; Fisher's z, the t, F, and χ^2 distributi simple cases of analysis of variance and of interaction; and the var

sures of the idea of correlation (correlation index and ratio as well as ficient), tetrachoric, biserial, partial, multiple, and rank correlation. All is done without the use of the calculus, and is usually developed in d pedagogic fashion. The emphasis on sampling effects is good, and reference to $n = 30$ as the boundary between 'small' and 'large' is il in connection with the usual size of a teacher's class.

ach of the fifteen chapters includes, in addition to the text, the necessary es and figures, exercises, and references. The references, whether by note or at the end of a chapter, are nearly all to American workers, gh a chart from *Biometrika* is taken via an American book. At the end nine appendices. Two of these are tables of data on which some of the cises are based, and one is a 10×10 set of blocks of 5×5 numbers n from the Rand Million Random Numbers. All the others are of dard statistical functions. There are very few misprints, the printing is and the binding sound. On the whole the expositions of the writer very clear, but there are a few cases where the meaning is not imme-ly obvious. The computations, judged by some sample checks, are ect.

though the references are nearly all from American books, yet British ers are fairly treated. It would have been gracious, I think, to have e reference to the 1888 paper of Edgeworth on the distribution of nination marks. The assumptions in connection with any ordinary e' as used in educational work do not really justify the inclusion of such le in any of his categories on page 17, while the definition of IQ, though opriate originally, is not valid for most of the common British group The difficulty of a repeated draw of the same school in the sampling mple given on page 119 ought to have been mentioned. The data of endix A should have been given serial numbers in order that the selec-asked for in Chapter 10, Exercise 2, could be conducted through in the ired manner.

conclusion, the book can on the whole be recommended not merely to ducational research worker, for whom it is primarily intended, but to others who want, in their use of the simpler statistical techniques, to l errors arising from some of the commoner unjustified assumptions.

FRANK SANDON

da: *The Theory of Games and Linear Programming* (Methuen's Monographs on Physical jects). London: Methuen and Co.; New York: John Wiley and Sons, 1956. Pp. 106. 6d.

is generally recognised that the subject of mathematical probability, as n today, and in direct line of descent the subject of mathematical tics, grew from the humble enquiries of gamblers centuries ago. Indeed, urse of lectures on probability theory is rarely given without some nce to the tossing of a die or to some card game. In the consideration mes of pure chance the mathematical analysis has so far been quite actory, but little success has been obtained in the analysis of more licated games, such as chess. A player in a game of chess is not ruled by the laws of chance, but may derive benefit from an intelligent line y, or strategy. It is the inclusion of this important element that has rise to a new sub-field of mathematics known as the 'theory of games ategy', on which topic a number of books, mostly American, have been shed in recent years.

Another subject of recent development, and at first sight unrelate game theory, is Linear Programming. In its most general form the pro here is the optimisation of a linear form subject to a number of linear straints, and the technique has been applied widely to economic prob in which profit is to be maximised or costs minimised.

The present volume, in the series of Methuen Monographs on Phy Subjects, provides a short account of game theory, demonstrates the between this and linear programming, and goes on to describe the algeb and graphical approaches to the latter. The very fact that all this is achie within 106 small pages is an indication that the treatment is always con but the author is to be congratulated on the fact that it is mostly clear readable. The book starts with a chapter outlining the theory of game which the necessary terminology, and the underlying minimax princi are described with the help of a number of examples. The second cha describes the graphical representation of game theory, and the next is dev to a proof of the 'fundamental theorem'. The following five chapters concerned with linear programming, and include discussions on its graph representation, the algebra of the simplex method, degeneracy, and dua In chapter nine a method of reducing any game to a linear programm problem is described, and several games are solved in this way. Fin there is a further discussion of the graphical representation of linear gramming, and a brief description of the alternative approach by E. M Beale known as the method of leading variables.

Other books on game theory have attempted to trace an association statistics and, in particular, with statistical inference. Fortunately no attempt is made in this book, the object of which (as stated on the cove that 'economists, mathematicians, statisticians, operational research wo and many others are offered easy access to a subject that should be of int and of value to them.' The main criticism of the content is that, from point of view of many of the people for whom this book is intended, t will seem to be a paucity of examples showing just how game theory linear programming may be of some use in practice.

This drawback apart, however, the book is an excellent, united acc of two new and interesting subjects, and its size makes it such that it ca read easily from cover to cover rather than treated as a book of refer

G. A. CO

EVOLUTIONARY OPERATION:*
IETHOD FOR INCREASING INDUSTRIAL PRODUCTIVITY

GEORGE E. P. BOX†
Imperial Chemical Industries Limited
Dyestuffs Division Headquarters, Manchester

The rate at which industrial processes are improved is limited by the present shortage of technical personnel. Dr Box describes a method of process improvement which supplements the more orthodox studies and is run in the normal course of production by plant personnel themselves. The basic philosophy is introduced that industrial processes should be run so as to generate not only product, but also information on how the product can be improved.

duction

Such scientific effort in industry is directed on the one hand to the overy of new products and processes and on the other to their lopment and improvement. This paper is concerned with a par-ar aspect of the problem of *improving* industrial processes.

dustrial organisations usually have specialist groups of scientific kers in research, development, and experimental departments, manently occupied with improving manufacturing processes, who loy a wide variety of techniques, ranging from the fundamental y of reaction mechanisms to the purely empirical assessment of effects of changes in variables. Associated experimentation may onducted in the laboratory, on the pilot plant, and on the full process, and in particular may involve the use of statistical tech-es having a fairly high degree of sophistication.[1–7] As a result of application of this variety of specialised effort a steady rate of ease in productivity is usually attained.

timately the rate of improvement is limited by the shortage of nical personnel. This shortage can be expected to become more er than less severe, and in searching for further ways of attaining ter process efficiency one must look for methods which are sparing eir use of scientific manpower. The object of this paper is to out-one such device, which has been applied with considerable success the past few years.

his is called 'Evolutionary Operation'. It is a method of process ation which has a 'built-in' procedure to increase productivity. es some simple statistical ideas and is run during normal routine

sed on a paper given at the International Conference on Statistical Quality Control sed by the European Productivity Agency of OEEC in Paris in July 1955.

ow Director of the Statistical Techniques Group at Princeton University.

production largely by plant personnel themselves. Its basic philosc is that it is nearly always inefficient to run an industrial proces produce product alone. A process should be run so as to gene product *plus information on how to improve the product.*

The technique is in no sense a substitute for the more fundame investigations referred to above. On the contrary, the effects covered by the application of evolutionary operation, particul those which are of an unexpected kind, help to indicate new a where fundamental research might be rewarding. Although method has been specifically developed as a production technique the chemical industry, it is believed that it has more general appli bility.

Plant-scale and Small-scale Experiments

In the chemical industry the plant process will usually have b arrived at after considerable experimentation on the small scale. N the optimum conditions of operation on the small scale usually prov no more than a good first approximation to the full-scale optim Because of this it is commonly found that considerable modificat of the conditions arrived at from small-scale work is necessary be a comparable result can be obtained on the plant itself.

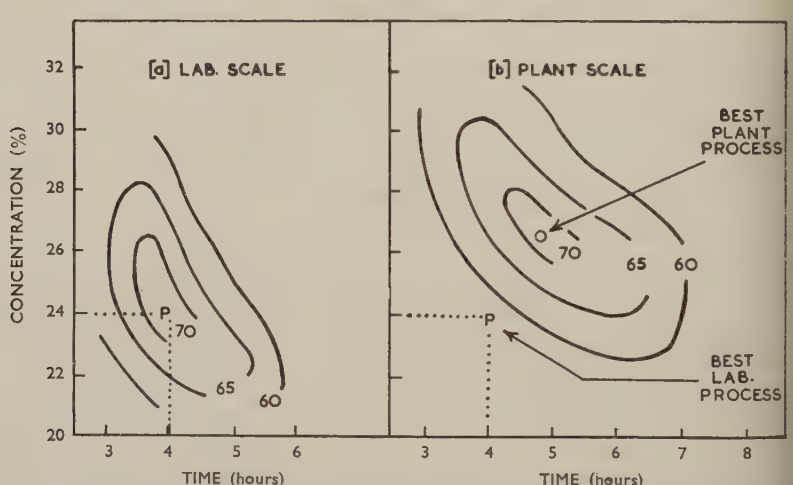


FIG. 1. Possible appearance of yield surfaces, showing contours of percentage yield, f process conducted (*a*) on the laboratory scale and (*b*) on the plant scale.

Fig. 1 shows how, in translating the process from the laboratory the plant, the yield surface can become distorted and displaced to scale-up effects. It will be seen that when this happens one mi expect the best laboratory conditions to give disappointingly low yi on the full scale. The effort to move from the laboratory maximu yield conditions at P to the plant maximum-yield conditions at O m evidently be exerted *on the plant itself* and not in the laborator investigation in the laboratory can only lead back to P.

or those unfamiliar with the representation adopted in Fig. 1 it ld be explained that the relationship between the response ('yield') the process variables ('time of reaction' and 'concentration of one ne reactants') is imagined to be represented by a solid graph or ponse surface'. In the neighbourhood of maximum yield such a ace may have the appearance of a mound. The height of the nd at any particular point represents the yield at some set of tion conditions. To allow representation in two dimensions the l is shown by contours in the same way that the height of land is esented by contours on a map. Although for simplicity the above ussion is conducted in terms of yield, it should be understood that ditions giving highest yield would often not represent the optimum ess. There would, for example, be no advantage in obtaining a er yield if to do so involved the use of a disproportionate amount me expensive starting material. The principal response usually idered, therefore, is 'the cost of producing unit quantity of product er the specified manufacturing conditions', or some other measure roductivity which takes account of the cost of running the process. addition to improvements made possible by adjustment of process ditions already studied on the small scale, further progress is usually ble by the introduction of new modifications not considered—and not capable of being studied—at the small-scale stage of develop- t.

djustments are made when the plant is first installed, but these m result in the location of the ultimate plant optimum, and as a t of special experimental campaigns, chance discoveries, and new , improvement usually continues over many years. The object of tionary operation is to speed up this process.

gy with Evolutionary Process

ne method used to speed the improvement is illustrated by the wing analogy. Living things advance by means of two mechanisms:

) Genetic variability due to various agencies such as mutation.

) Natural selection.

nical processes advance in a similar manner. Discovery of a new for manufacture corresponds to a mutation. Adjustment of the ess variables to their best levels, once the route is agreed, involves ocess of natural selection in which unpromising combinations of evels of process variables are neglected in favour of promising ones. g. 2 illustrates diagrammatically the possible evolution of a species bster. It is supposed that a particular mutation produces a of lobster with 'length of claws' and 'pressure attainable between ' corresponding to the point P on the diagram and that in a environment the contours of 'percentage surviving long enough produce' are like those shown in the figure.

The dots around P indicate offspring produced by the initial
of lobster. Since those in the direction of the arrow have the gre
chance of survival, over a period of time the scatter of points re
senting succeeding generations of lobsters will automatically move up the survival surface. This automatic process of natural selection ensures, without any special effort on the part of the lobsters, that optimum-type lobsters exist. It also ensures that if the environment changes so that the survival surface is altered, the lobsters will change correspondingly to the new point of maximum survival.

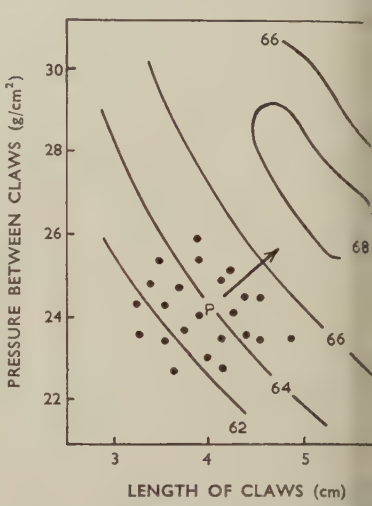


FIG. 2. Evolution of a species of lob
Contours show the percentage surv
long enough to reproduce in a give
vironment.

What we have to do is to imitate this process. That is to say, we have to institute a set of rules for *normal plant operation* so that (without serious danger of loss through manufacture of unsatisfactory material) an evolutionary force is at work which steadily and automatically moves the process towards its optimum condit
if it is not operating there already. Such a technique will gradu
nudge the operating procedure into the form that is ideally suite
the particular piece of equipment which happens to be available.
two essential features of the evolutionary process are:

(i) *Variation.*

(ii) *Selection* of 'favourable' variants.

Static and Evolutionary Operation

Routine production is normally conducted by running the plan
rigidly defined operating conditions called the 'works process'.
works process embodies the best conditions of operation known at
time. The manufacturing procedure, in which the plant operator a
always to reproduce exactly this same set of conditions, will be ca
the method of *Static Operation.* Although this method of operation
strictly adhered to, clearly precludes the possibility of evolution
development, yet the *objectives* which it sets out to achieve are ne
theless essential to successful manufacture, for in practice we
interested not only in the productivity of the process, but also in
physical properties of the product which is manufactured. Th
physical properties might fall outside specification limits if arbitr
deviations from the works process were allowed. Our modified met
of operation must therefore include safeguards which will ensure t

isk of producing appreciable amounts of material of unsatisfactory
ty is acceptably small.

the method of Evolutionary Operation a carefully planned cycle
nor variants on the works process is agreed. The routine of plant
ation then consists of running each of the variants in turn and
nually repeating the cycle. The cycle of variants follows a simple
rn, the persistent repetition of which allows evidence concerning
ield and physical properties of the product in the immediate
ity of the works process to accumulate during routine manufac-
In this way we use routine manufacture to generate not only the
uct we require but also the information we need to improve it.

ntrolled variation having thus been introduced into the manu-
re, the effect of selection is introduced by arranging that the
ts are continuously presented to the plant manager in a way which
ily comprehended. This allows him to see what changes ought
made to improve manufacture. The stream of information con-
ng the products from the various manufacturing conditions is
narised on an *Information Board* prominently displayed in the plant
ager's office. This is continuously brought up to date by a clerk
om the duty is specifically assigned. The information is set out
ch a way that the plant manager can at any time see what weight
idence exists for moving the centre of the scheme of variants to
new point, what types of change are undesirable from the stand-
of producing material of inferior quality, how much the scheme
ting to run, and so on.

making a permanent change in the routine of plant operation
ituation is very different from that which we meet in running
alised experiments on the plant. The latter will last a limited
during which special facilities can be made available. Further-
some manufacture of substandard material is to be expected and
be budgeted for. Evolutionary operation, however, is virtually a
nent method of running the plant and cannot therefore demand
al facilities and concessions. For this reason only small changes
e levels of the variables can be permitted, and only techniques
e enough to be run continuously by works personnel themselves
r actual conditions of manufacture can be employed.

e effects of the deliberate changes in the variables will usually be
ed by large errors customarily found on the full scale. *However,*
roduction will continue anyway, a cycle of variants which does not signifi-
effect production can be run almost indefinitely, and because of constant
ion the effect of small changes can be detected.

xample

illustrate the procedure we consider one phase of evolutionary
tion for a particular batch process. At this stage of development
rocess factors—the percentage concentration of one of the reac-
and the temperature at which the reaction was conducted—

were being studied following the scheme of variants shown in F
The works process is labelled (1) and the four variants are labelle
(3), (4), and (5). One batch of product was made at each set of
ditions, which were run successively in the order 1, 2, 3, 4, 5; 1,
4, 5; and so on. Three responses were recorded:

(i) The cost of manufacturing unit weight of product. This was obtained by dividing 'the cost of running at the specified conditions' by 'the observed weight yield at those conditions'. It was desired to bring this cost to the smallest value possible, subject to certain restrictions listed in (ii) and (iii) below.

(ii) The percentage of a certain impurity. It was desired that this should not exceed 0·5.

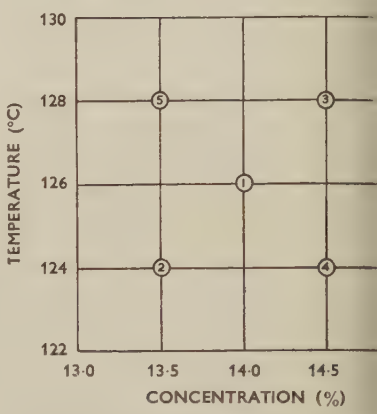


FIG. 3. A cycle of variants about the v process.

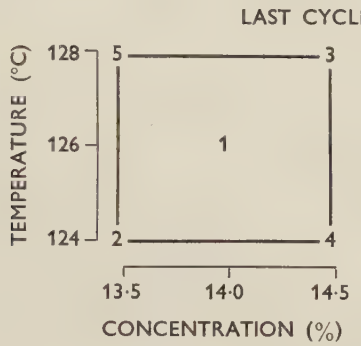
(iii) A measure of fluidity. It was preferred that this should between the limits 55 and 80.

The information coming from the experiment was recorded by wri in chalk on an ordinary blackboard. Alternatively, wax pencil white plastic board or magnetic letters and numbers on a steel b could have been used. The essential thing is that it should be a sir matter to erase or remove one number and replace it by another. scheme set out in Fig. 4 is not the only one which could have adopted, but is intended to show the sort of calculations and la of the results which have been found useful.

The phase number at the top left-hand corner of the board indi that two previous phases of evolutionary operation have alr been completed on this process. In general these might have invo other variables or the same variables at other levels. In order the new results may be considered in proper relation to those obta previously, the final average values recorded in previous phases sh also be available (for example on sheets of paper pinned to the boa The cycle number in the top right-hand corner indicates that 16 cy of this third phase of operation have been completed. There foll a plan of the cycle of variants being run.

The table below this summarises the current situation. First shown the requirements which it is desired to satisfy. These are follo by the running (i.e. up-to-date) averages at the various manufactu conditions set out so as to follow the plan of the cycle of variants. arrangement makes it easy to appreciate the general implications o results. A measure of the reliability of the individual running aver is supplied by the '95% error limits'. These are simply the quant

$\sqrt{n}$ appropriate to the calculation of fixed sample-size confidence s for individual means. In this expression s is the standard deviation n the number of cycles, and t the appropriate significance point of ent's t distribution. The scheme can be run until these limits for neans of the principal response have been reduced to an acceptably l preassigned width.[8]



		Cost	Impurity (%)	Fluidity
Requirement		Minimum	Less than 0·50	Between 55 and 80
Running Averages		32·6 33·9 32·8 32·3 33·4	0·29 0·35 0·27 0·17 0·19	73·2 76·2 71·3 60·2 67·6
5% Error Limits		±0·7	±0·03	±1·1
ffects with 5% Error Limits	Conc.	1·2 ± 0·7	0·04 ± 0·03	5·2 ± 1·1
	Temp.	0·4 ± 0·7	0·14 ± 0·03	10·8 ± 1·1
	C × T	0·1 ± 0·7	0·02 ± 0·03	−2·2 ± 1·1
	Change in Mean	0·2 ± 0·6	−0·02 ± 0·03	−1·6 ± 1·0
Standard Deviation		1·44	0·059	2·12
95% Error Limits		1·22 1·76	0·050 0·072	1·80 2·59
Prior Estimate		2·71	0·054	3·22

Fig. 4. Appearance of information board at the end of cycle 16.

low this are shown the 'effects' of the variables and their 95% limits. The concentration and temperature effects are each calculated in the usual way as the difference in the average values of the onse at the higher and the lower level of the variable. The value e centre conditions does not enter the calculations except in computing the 'change in mean' effect. If y_1, y_2, y_3, y_4, and y_5 are the ing averages of one of the responses after n cycles of operation, ffects and their limits of error for this particular example are as ws:

Effect	*Value*	*Limits of Err*
Concentration	$(y_3 + y_4 - y_2 - y_5)/2$	$\pm ts/\sqrt{n}$
Temperature	$(y_3 + y_5 - y_2 - y_4)/2$	$\pm ts/\sqrt{n}$
Interaction	$(y_2 + y_3 - y_4 - y_5)/2$	$\pm ts/\sqrt{n}$
Change in mean	$(y_2 + y_3 + y_4 + y_5 - 4y_1)/5$	$\pm 2ts/\sqrt{(5}$

The effect referred to as the 'change in mean' is simply the grand n for all the runs $(y_1 + y_2 + y_3 + y_4 + y_5)/5$ less the mean for 'standard' conditions y_1. It is therefore an estimate of the differenc average response resulting from the use of the evolutionary sche In the present example the effect of introducing the evolutior scheme is thus:

(i) To increase the cost per batch by $0{\cdot}2 \pm 0{\cdot}6$ units.

(ii) To reduce the average impurity by $0{\cdot}02 \pm 0{\cdot}03\%$.

(iii) To reduce the average fluidity by $1{\cdot}6 \pm 1{\cdot}0$ units.

(It will be seen that if the effect of blending batches of slightly differ qualities was to average the physical properties, the 'change in me would measure the difference between the product of the works pro and a blend of the products from evolutionary operation. In m actual examples partial or complete blending of the products c naturally occur at the later stages of manufacture, so that wl physical properties behave approximately additively there may in be remarkably little overall change in the manufactured product to evolutionary operation. In some cases, especially if the effec introducing small variations in the levels of the variables was u pectedly large, blending could be deliberately introduced to proc an acceptable product.)

If the variants cover a region of the response surface which sloping plane, then it is not difficult to see that the true 'chang mean' will be zero. For a convex surface such as that near a m mum it will be negative, whereas for a concave surface such as near a minimum it will be positive. It can be shown that, on cer plausible assumptions, the 'change in mean' is proportional to the of the quadratic constants which measure curvature in the direct of the variables. When the interaction effect and the change in m effect are not small compared with the single effect of the variab this indicates that a maximum or minimum is being approached and exact location and exploration a technique fully set out elsewher is adopted.

For the cost response the change in mean supplies a continu measure of the cost of obtaining information by the process of ev tionary operation. In the present example this is estimated at $0{\cdot}2 \pm$ units of cost per batch. If the cost surface were locally planar, ev tionary operation would cost nothing. In practice, concavity of cost surface is to be expected, since a minimum is usually b approached, so that there will usually be a small cost associated v

ning the evolutionary process. Except when the process has been ught very close to its ultimate optimum this cost will be redeemed ny times over by the value of the permanent process improvements occur from time to time as a result of the information generated he evolutionary scheme.

fter the calculated effects, the experimental error standard devia- s calculated from the observations themselves are shown. Except he initial stages of the scheme these are used in computing the ts of error for the running means and for the 'effects' of the ables. The normal theory fixed sample-size 95% confidence limits hese estimates of the standard deviations are also shown. The final ns are the estimates obtained from prior data used in initiating the eme.

n general, by inspecting the results set out on the information board he light of the requirements it is desired to satisfy and from expert wledge of other factors which affect plant operation, the plant nager decides at any particular stage whether

a) to wait for further information;

b) to modify operation.

der (*b*) some of the alternatives open to him are:

i) To adopt one of the variants as the new 'works process' and to recommence the cycle about this new centre point.

ii) To explore an indicated favourable direction of advance and recommence the cycle around the best conditions found. (This exploration may be done, for example, by making a series of tentative advances in the indicated direction, at each stage running the new conditions and the previous best conditions alternately, until sufficient evidence has been gathered.)

i) To change the pattern of variants to one in which the levels are more widely spaced.

v) To substitute new variables for one or more of the old variables.

In the example actually discussed it is seen that a decrease in concentration would be expected to result in reduced cost, reduced impurity, and reduced fluidity. The effect on cost of a decrease in temperature is uncertain but is more likely to be favourable than not, and would almost certainly result in marked reductions in impurity and fluidity. These facts have to be considered, bearing in mind that a fluidity of less than 55 is undesirable and that although a further large reduction in impurity is welcome it is not necessary in order to meet the specification. It was decided in the event to explore the effects of reducing concentration alone. Phase 3 was terminated, and in the next phase the three processes (13%, 126°), (13·5%, 126°), (14%, 126°) were compared. The first of these gave a mean cost of 32·1 with an impurity of 0·25 and a fluidity of 60·7, and was adopted as a base for further development.

A geometrical display of the results on the information board m also be used when three variables are jointly considered. In this c perspective drawings of the sort illustrated in Fig. 5 are used. T

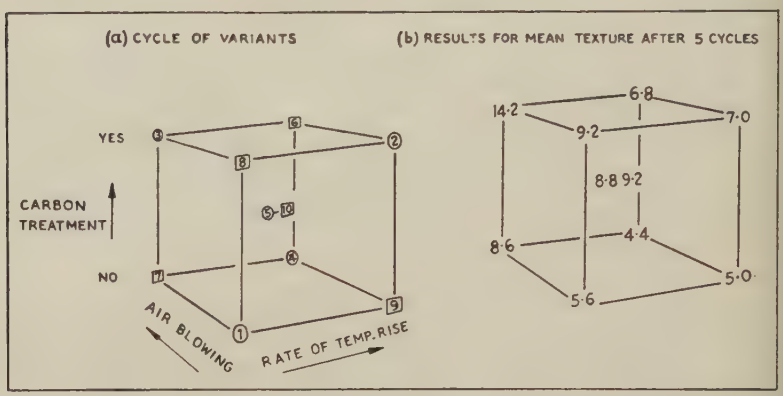


FIG. 5. Display of pattern of variants and results for three variables.

cycle of variants for a three-variable scheme is shown in Fig. 5(*a*) while 5(*b*) shows the results from one of the responses after five cycl of operation. This method may, of course, be applied to respons which cannot be measured directly. In the case illustrated in Fig. 5(an important property of the product was a somewhat esoteric quali referred to as 'texture'. A set of artificial standard samples were pr pared which were judged by experts to have a range of 'textures' approximately uniform steps, and these were arbitrarily scored. T texture of a sample of each manufactured material was then match against the standards and an appropriate score given to it. In a simil way a scheme to evolve conditions which would give a product le inclined to 'cake' has been run, in which caking was judged by visu inspection and scored by comparison with a verbally defined scale.

Selection of the Variants

The technique outlined differs from the natural evolutionary proc in one vital respect. In nature the variants occur spontaneously, b in our artificial evolutionary process we have to introduce the Variants involving the levels of temperature, concentration, pressur etc., are natural choices, but there are usually an almost unlimite number of less obvious ways in which manufacturing procedure ca be tentatively modified. Frequent instances of marked improvemer due to some innovation never previously considered in a process whic has been running for many years testify to the existence of valuab modifications waiting to be thought of.

To make our artificial evolutionary process really effective, therefor one more circumstance is needed—we must set up a situation in whic useful ideas are continually forthcoming. An atmosphere for th generation of such ideas is perhaps best induced by bringing togeth

suitable intervals a group of people with special, but different, nnical backgrounds. In addition to plant personnel themselves, rious candidates for such a group are, for example, a research man h an intimate knowledge of the chemistry of the similar reactions l a chemical engineer with special knowledge of the type of plant question. The intention should be to have complementary rather n common disciplines represented.

These people should form the nucleus of a small evolutionary opera- n committee, meeting perhaps once a month, whose duty it is to p and advise the plant manager in the performance of evolutionary ration. The major task of such a group is to discuss the implications current results and make suggestions for future phases of operation. eir deliberations will frequently lead to the formulation of theories ch in turn suggest new modifications that can be tried with profit. ince questions of modification of certain physical properties of the nufactured product may arise, a representative of the department ponsible for the quality of manufacture should also be on the evo- onary operation committee. Rather more may be got from the ılts and more ambitious techniques adopted if a statistician is also sent at the meetings.

With the establishment of this committee all the requirements for efficient evolutionary method of production are satisfied and the sed loop' illustrated in Fig. 6 is obtained. We are thus provided h a practical method of 'automatic optimisation' which requires special equipment and which can be applied to almost any manu- uring process, whether the plant concerned is simple or elaborate.

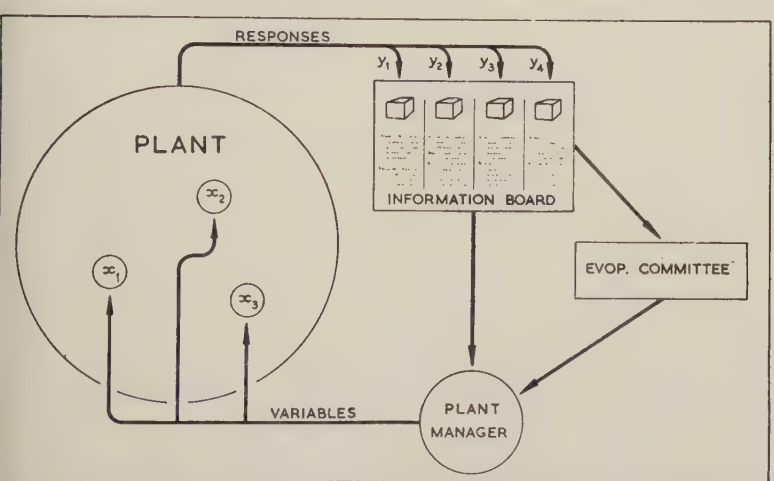


Fig. 6. Diagrammatic representation of the 'closed loop' provided by Evolutionary Operation.

t the beginning of this article I spoke of evolutionary operation as g run largely by plant personnel themselves rather than by spe- sts. The use of some specialists as advisers on the evolutionary

operation committee does not seriously vitiate this principle. In p
tice, the time spent by the specialists is perhaps one afternoon a mo
and the ultimate responsibility for running the scheme still rests
the plant manager and not with the specialists.

When not to Stop

With an alert team of workers new ideas should be continually fo
coming and the evolutionary method becomes virtually a perman
mode of operation and should be so regarded. Only if it seemed t
more would be lost than gained from the evolutionary proced
would the reintroduction of static operation be justified. In prac
it is found that even very small gains will justify the continual operat
of the evolutionary method. The situation at any given time can
appraised by the use of a pictorial log like that shown in Fig. 7.

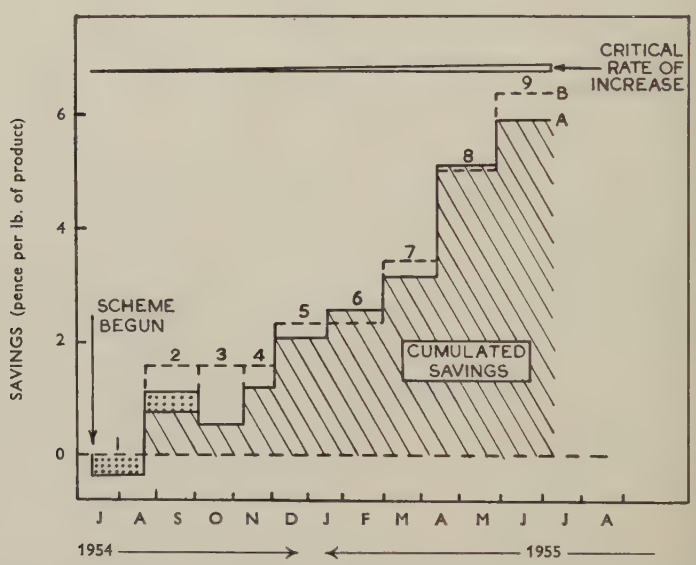


FIG. 7. Progress of an evolutionary scheme, showing critical rate of increase.

The full line A in this diagram shows the savings in pence per
which have been achieved in the various phases of operation. T
dotted line B shows the savings that would have resulted if the centre
'works process conditions', appropriate for each particular phase, h
been run. On the assumption of constant output the shaded area
proportional to the accumulated savings resulting from the schem
while each of the rectangular areas between the dotted and full li
shows the accumulated 'expenses' of running the scheme during th
phase. The speckled area in phase 1 represents the cumulated expen
for the scheme during this phase and is debited from the cumulat
savings in phase 2.

Whereas each phase of evolutionary operation, and consequently the nses associated with it, lasts for only a limited time, any improve- ts which result go on for as long as the process is used. Suppose assumed that process improvements will go on earning money for ars after they are discovered and that the running of the evolu- ary scheme adds c pence per lb. to the cost of the product. Then question of whether or not, at any instant of time, evolutionary ation should be continued may be resolved by comparing the rate nprovement r which it is expected may be produced by the evolu- ary process (measured in pence per lb. per year) with the critical of improvement r_0 given by $r_0 = c/p$. For if it is expected that the utionary scheme will need to be run for time t years to produce mprovement at the end of that period of rt pence per lb., and if k f product is made per year, then the total saving during the p years vhich the discovery is used will be $rtkp$ pence. During this time, . of product will be made and the loss due to running the evolu- ary scheme will be ckt pence. Thus the scheme will pay off if $rtkp$ eater than ckt, that is if r is greater than $r_0 = c/p$.

s an example consider the situation in Fig. 7. Should the scheme e shown be continued or not? Let us suppose (very conservatively) improvements on this process are expected to go on earning money 3 years after their discovery, so that p is put equal to 3. Suppose that c is taken to be the average of the values experienced in the evious phases. This gives the value $c = 0{\cdot}3$ pence per lb. We then for the critical rate $r_0 = 0{\cdot}3/3 = 0{\cdot}1$ pence per lb. per year. Thus ng as the rate of improvement due to the evolutionary process is cted to be at least as great as 0·1 pence per lb. per year the evolu- ary scheme should be continued. This critical rate of increase is vn diagrammatically at the top of Fig. 7. It will be seen that the al rate of improvement which had been experienced over the pre- s year was about 6 pence per lb. per year (about 60 times the cal rate), and there is no evidence as yet of any flagging in this of improvement. There is therefore no doubt whatever that this me should be continued.

he example given is by no means atypical, which explains my tence that the evolutionary method should be regarded as virtually rmanent mode of operation. It is psychologically wrong to talk roduction under such a scheme as 'experimental manufacture', an experiment is something which is done for a limited period is not part of the normal run of things.

Questions and Tentative Answers

ke statistical quality control, evolutionary operation is designed for cation in the factory itself. Its aims are different and more ambi- than those of quality control, however, since it is directed to re not a more uniform product but a cheaper and better product. ieve that this technique, if applied sufficiently widely, can have a

marked affect in achieving greater productivity in industry by ensu that the plant that is available, whether old or new, is operated in best possible manner. The outline above is intentionally gen because it is the general attitude and philosophy that is impor here, and not the particular details of application. A full accou nevertheless being prepared in which a number of technical quest are discussed. In the present paper I shall do no more than indi some of these questions and some tentative replies, which it is hc to amplify and to justify in the later discussion.

Q. How many variables should be included at one time in an ev tionary scheme?

A. Usually two or three variables can be handled satisfactorily un the normal production conditions with which I am familiar. It sho perhaps be emphasised once again that what is being discussed is normal production situation in which evolutionary operation is plied. For specialist short-period investigations (which, as has b explained in the introduction, are *not* the subject of this account which nevertheless play an extremely important part in the gen scheme of process development to which evolutionary operation supplement) the situation is entirely different. In these specialist vestigations where, *for a limited period*, it is permissible to interfere v production, and where special supervision and other facilities are m available, the object should be to saturate (or following Satterthwa possibly 'super-saturate') the experiment with as many factor possible.

By studying the variables in groups of three or so at a time we fo the possibility of detecting dependence (interaction) between varia not in the same group. The effect of this limitation should be m mised as much as possible by examining in the same group se variables which are expected to be interrelated. Periodically varia whose effects have been found to be important in different phas operation should be tested together.

Q. What patterns of variants are of most value?

A. A variety of patterns of variants are useful for particular purp Among the most valuable for initial use are those based on two-l factorial designs with one or more added points at the centre conditi These are the arrangements shown for two and three variables in F 3 and 5. They have the advantages that:

(i) They are simple to comprehend, perform and analyse.

(ii) The added centre points allow continual reference to the 'st dard' process and permit the 'cost' of the evolutionary scheme to assessed.

(iii) Complexity of the surface is easily detected by considering relative magnitudes of the simple 'main effects' on the one hand the 'change in mean' and interaction effects on the other.

) They can be made the nucleus of more elaborate designs (in rticular of composite second-order rotatable designs[10]) by which mplexity of the surface may be elucidated.

They lend themselves conveniently to 'blocking arrangements' ereby extraneous disturbances due to such uncontrolled factors as ne trends may be reduced. In Fig. 5(*a*), for example, the circles d squares indicate two sub-cycles into which the complete cycle ay be divided. A general change in mean occurring between sub-cles will not bias the estimation of effects.

How should past plant-records be used in planning the evolu-ry scheme?

When, as is often the case, past plant-records covering long ds of normal operation are available, the planning of an evolu-ry scheme should always begin with a careful study of these ds. In particular they may be used to determine the approximate nitude and nature of the uncontrolled variation in the various nses, and consequently the number of repetitions of the cycle to be needed to detect effects of a given size.

om these records we can find out whether the errors in the prin- response can be regarded as effectively independent and, if not, an determine the nature of the dependence that exists. This is of importance in choosing the period for which each variant should un before changing to the next variant. In practice, of course, period depends partly on convenience of operation and, for con-us processes, on the time it takes for the plant to settle down a change in reaction conditions. The time for running each nt which gives the maximum amount of information for a given period of production can be shown to depend on the nature of ependence between the observations, and may be determined by rly simple use of the correlogram along the lines considered by tt[11] or, equivalently, by considering the spectral properties of the d.

Should the variants be run in random order?

Faced with the possibility of serial correlation between successive vations, the statistician would normally wish to perform the vari-n random order within each cycle, thus guaranteeing the validity simple type of analysis used in the example of Fig. 4. However, me cases, particularly where the time for running each cycle is it is much simpler to run a systematic routine of variants on the than a random one. In these circumstances randomisation is ly abandoned (as was in fact done in the example considered). bservations after n cycles of k variants can be written in a table g n rows and k columns, and we are only concerned with com-n of *column* means. Now the major part of the dependence occurs rows, and in this situation, as was pointed out by R. A. Fisher,[14,15] mple analysis of the type we have considered above will usually

not be seriously invalidated. Where the correlogram is available possible to determine how far dependence between observations affect the simple analysis and what remedial measures, if any, nee taken, although this refinement would seldom be worth while.

Q. How should multiple responses be considered?

A. Although it is theoretically possible to equate all responses single criterion such as profitability, this usually presents great prac difficulties. As a general rule it is best to represent the problem as of improving a *principal* response (for example the cost per lb product) subject to satisfying certain conditions on a numbe *auxiliary* responses. These auxiliary responses usually measure quality and important physical properties of the product.

Very careful thought in the selection of the principal respons essential. The vital question to ask is: 'If this response is improved it mean *necessarily* that the process is improved?'

In the example of Fig. 4 the reconciliation of the requirements the various responses in the light of the experimental results was d intuitively and led to the decision to reduce concentration alone. intuitive approach has the virtue of simplicity and allows backgro information not emanating from the experimental results themselv be taken into account. It is fairly satisfactory in the situation spe cally dealt with here when there are only two or three variable consider. However, as has been pointed out elsewhere,[12] the prob is really one of programming, in the sense of linear programming, the added complications that the problems are not always app mately linear and that the restraints are not known exactly but r be estimated. In the fuller account we show how certain calcula can help with the more difficult cases.

Q. How best can the stream of information coming from the p during the evolutionary process be presented to those responsibl deciding what to do?

A. Two things are necessary: first, to show how much weight o to be attached to the results, and second, to present them in such a that their interpretation is facilitated as much as possible.

To convey a sense of the degree of reliability which the plant man should associate with the results, a number of ideas have been tr In the original schemes various types of sequential charts and sig cance tests were used. It is now felt, however, that the problem is one of significance testing and that what is needed is a presentatio the information contained in the data unweighted by external tures subsequently injected into the situation. For example, the p ticular choice of the risks α and β and of the hypotheses 'whic is desired to test' (subtleties not readily comprehended by pl managers) can completely alter the apparent implications of a se data when these are plotted on a sequential chart. If the observat are roughly normally distributed, are independent, and have cons

nce, then all the information they contain is included in the mean, tandard deviation, and the number of observations. These statis- eem best comprehended in the form of a mean with its 95% con- ce limits. In appraising the results prior information about the rtance of different sorts of effect must be used. It seems best, ver, to separate this from the presentation of the results, which refer to information supplied by the observations and to nothing This problem is regarded as being far from solved. It involves intricacies which cannot be discussed in the present account and obably best considered in terms of stochastic approximation[16] and -mechanism theory. All that would be claimed for the present od is that it does allow a satisfactory evolutionary process to go on. show the implications of the mean results once it becomes rent that these are determined sufficiently accurately, there is no t that for two or three variables geometrical representation, such at shown in Figs. 4 and 5, is ideal. It allows the general trend e responses and their relationship to each other to be appreciated manner not possible in any other way.

e plant manager should run evolutionary operation in much the way as he would play a card game. The information board s him his 'hand' at any given time, and depending on that hand are a number of actions he can take (including drawing a further and deciding what to do then).

here one or more of the variants is clearly better than the works ss or where clear-cut trends in the results exist, the plant manager have no difficulty in following the indications of the information l. Where the results indicate that complexity exists he will be o obtain the help of the statistician on the evolutionary operation ittee in elucidating the results and, where necessary, in augment- nd modifying[4-6] the cycle of variants in the next phase of opera- o as to resolve the complexity.

duplicate information board may be kept on the plant itself and gnificance explained to process workers. This provides added st and is an incentive to accurate operation, which itself can in general improvement in productivity.

In what way can the results from small-scale experimental s be used in planning the evolutionary scheme, and how should ffect the way in which these small studies are conducted?

The complexity mentioned above arises principally because the bles studied fail to behave independently in their effects on the nse; that is, they interact. The plant process will usually have arrived at as the result of a small-scale investigation of at least of the variables. This investigation should have culminated in ly of the local 'geography' of the response surfaces in the neigh- ood of the proposed operating conditions. The principal features laboratory response surfaces will normally be preserved on the

plant scale even though some distortion occurs. If the characte of the laboratory surfaces have been determined in the manner tioned above, it is frequently possible to discover transforms o variables originally considered which act approximately independe at least for the principal response. By working in terms of these variables, difficulties due to complexity of the surface can be gr reduced in the plant-scale investigation.

Q. In practice it is impossible to attain truly static operation. S variations in the process conditions are bound to occur from one to the next. In cases where these changes are recorded, why sh one bother to carry out a special pattern of variants? Why not the 'pattern of variants' supplied by the natural variation of the pr to supply information on which evolutionary improvement can based?

A. It is true, of course, that for data generated by natural varia the simple type of analysis of the results which has been used a would no longer be applicable. However, this itself is no reason the natural pattern of variants should not be used. Suppose that level of response is denoted by y and that there are k variables w levels are denoted by $x_1, x_2, \ldots, x_k$; suppose that the works proc defined by the particular set of conditions $x_{10}, x_{20}, \ldots, x_{k0}$ and owing to imperfect control, fluctuations about those levels occur are recorded. Then we can, for example, assume a local relatio of the form

$$y = b_0 + b_1x_1 + b_2x_2 + \ldots + b_kx_k$$

and estimate the coefficients continuously by the method of squares (multiple regression). If our assumptions were correct, coefficients would measure the individual effects of the variables. calculations required to fit the equations are laborious but, as been suggested by Professor Goodman,[13] could in principle be mechanically (e.g. by an electronic computer).

At first sight such a method appears attractive, for here we see have an evolutionary scheme in which we do not need to bother a introducing variants deliberately. On closer examination, howeve value seems much more doubtful. Many investigations have made by statisticians over the years in which plant records have analysed by multiple regression in an attempt to determine the 'eff of the variables and so to improve the process. In my experience results of such investigations are nearly always disappointing. reasons are not far to seek:

1. Many of the factors that may vitally affect the efficiency of process are not in the normal course of events altered at all.

2. Those factors which vary naturally do so, not over the range should like, but over ranges dictated by the degree of control w happens to exist. The more control is improved, the less informa we get.

The fluctuations that naturally occur in the variables are often ly correlated. This results in poor precision of the estimates when y to disentangle the effects of the variables one from the other.

Accidental modifications often tend to happen in 'phases' and come spuriously correlated with causally unrelated time-trends sponse. Such effects can lead to completely wrong conclusions. npts to eliminate time-trends computationally usually eliminate ffects of the factors at the same time.

nat this all amounts to is that a naturally occurring scheme of nts is not very likely to provide a good, or even passable, 'design' consequently that the amount of information generated by natural tion may be scarcely worth salvaging.

Can evolutionary operation be made automatic?

The procedure of evolutionary operation so far described is a ual' one. It requires no special facilities and can be immediately ed in one form or another to a very large proportion of industrial sses. This is so whether the available plant is of the crudest kind nether it includes such refinements as automatic controllers and ders. The plant manager is himself a part of the 'closed loop', ensuring that sensible action will be taken even in unforeseen mstances.

th a sufficiently instrumented plant the evolutionary procedure course, capable of being made completely automatic. Thus bles whose levels are regulated by a controller can be automati- changed at regular intervals so as to follow a cycle of variants, a response such as cost per lb. can be automatically computed the readings of instruments which measure the properties of the uct. The cumulated differences in response at the various process tions can be used to trigger off adjustments in the location of the rn of variants, so completing the evolutionary process.

continuous processes (where there is a continuous input of starting rials and a continuous output of product) the 'pattern of variants', d of being a discrete set of points as in Fig. 3, can consist of a nuous locus. The problem of detecting the effects of the variables n precisely that which arises in communication theory, of detecting al of known form in a noisy channel.

e introduction of automatic evolutionary operation would usually rth while only if the response surface itself was changing in some nd it was desirable to attempt to follow that change. For many ical processes the response surfaces are reasonably stable. In however, unpredictable but steady changes can occur owing to changes in raw material (such as crude oil) or in catalyst activity. unpredictable differences in the position of the optimum condi- may occur between batches of catalyst and also within the life of le catalyst batch. In these cases automatic evolutionary operation

may be effective in keeping the plant operating near its best pe mance, but only if the rate at which information is generated is ciently large compared with the rate at which the optimum condi are changing. This is essentially a problem in the theory of se mechanisms.

Discussion

The device I have described is of course only a more powerful concentrated form of the naturally occurring evolutionary pro which goes on during all manufacture. In the ordinary cours events once the favourable effect of a deliberate or accidental modif tion is *recognised*, that modification is included in the works proc Unfortunately, because of a high level of variation, which usu obscures all but very large effects, favourable modifications freque go unrecognised unless they are forced to reveal themselves by device used here of constant repetition and consequent averaging- of errors.

That many of the problems touched on in the later part of paper are still the subject of active investigation should not obsc the fact that evolutionary operation, as set out in earlier sections, practical and immediately available method which ought to be m widely applied.

Both practical experience and theoretical consideration show very little can be lost and a great deal can be gained by applica of the technique, and for this reason evolutionary operation shoul adopted as a *normal production method*. Static operation shoulc tolerated only if good reasons for not using the evolutionary proce can be advanced.

REFERENCES

[1] Youden, W. J. (1954 onwards). Bimonthly articles on 'Statistical Design', *Indu and Engineering Chemistry*.

[2] Daniel, C. and Riblett, E. W. (1954). 'A multifactor experiment', *Industria Engineering Chemistry*, **46**, 1465.

[3] Vaurio, V. W. and Daniel, C. (1954). 'Evaluation of several sets of const and several sources of variability', *Chemical Engineering Progress*, **50**, 81.

[4] Box, G. E. P., Connor, L. R., Cousins, W. R., Davies, O. L. (Editor), H worth, F. R., and Sillito, G. P. (1954). *The Design and Analysis of In trial Experiments*. Oliver and Boyd, Edinburgh and London.

[5] Box, G. E. P. and Wilson, K. B. (1951). 'On the experimental attainmen optimum conditions', *J. R. Statist. Soc.*, B, **13**, 1.

[6] Box, G. E. P. (1954). 'The exploration and exploitation of response surfa some general considerations and examples', *Biometrics*, **10**, 16.

[7] Box, G. E. P. and Youle, P. V. (1955). 'The exploration and exploitatio response surfaces: an example of the link between the fitted surface the basic mechanism of the system', *Biometrics*, **11**, 287.

[8] Anscombe, F. J. (1954). 'Fixed-sample-size analysis of sequential observati *Biometrics*, **10**, 89.

TERTHWAITE, F. E. (1956). (Unpublished communication.)
, G. E. P. and HUNTER, J. S. (1956). 'Multifactor experimental designs for exploring response surfaces', *Ann. Math. Statist.* (In the press.)
ETT, G. H. (1955). 'The comparison of means of industrial time series', *Applied Statistics*, **4**, 32.
, G. E. P. (1955). Discussion at Symposium on Linear Programming, *J. R. Statist. Soc.*, B, **17**, 198.
KES, M. V. (1956). Discussion of paper 'Application of digital computers in the exploration of functional relationships' by G. E. P. Box and G. A. COUTIE at the Convention on Digital Computer Techniques. *Proceedings of the Institution of Electrical Engineers*, **103**, Part B, Supplement No. 1, 108.
ER, R. A. (1941). *Statistical Methods for Research Workers*, 8th edition, p. 226. Oliver and Boyd, Edinburgh and London.
, G. E. P. (1954). 'Some theorems on quadratic forms applied in the study of analysis of variance problems. II: Effects of inequality of variance and of correlation between errors in the two-way classification'. *Ann. Math. Statist.*, **25**, 484.
BINS, H. and MONRO, S. (1951). 'A stochastic approximation method', *Ann. Math. Statist.*, **22**, 400.

SOME STATISTICAL APPLICATIONS OF ELECTRONIC COMPUTERS

STEPHEN LIPTON

Rothamsted Experimental Station

In this article Mr Lipton describes in general terms son the work that has been carried out by the Statistics Departr at Rothamsted Experimental Station, using an electr computer. He pays particular attention to the routines w have been evolved in connection with the computational w

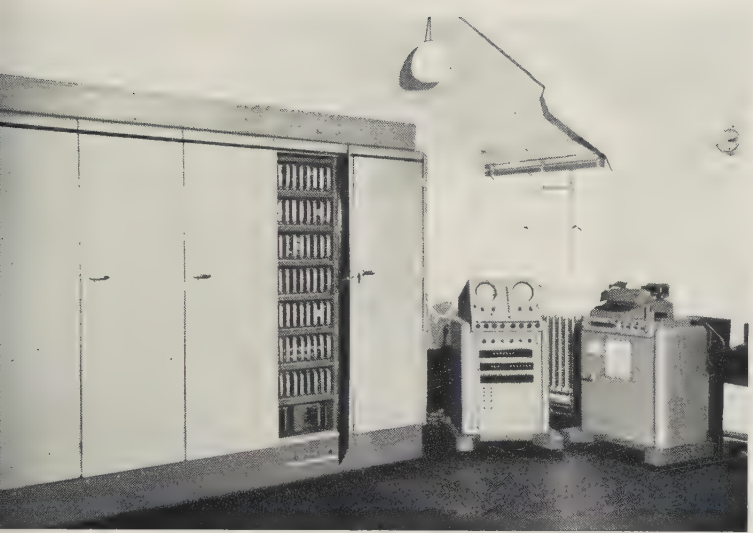
Introduction

The Computer at Rothamsted. The electronic computer installed in Statistics Department at Rothamsted is the Elliott 401 prototype mo This machine was built by Elliott Brothers under a development c tract with the National Research Development Corporation and i present being used mainly for statistical work in agricultural and logical research. A fairly detailed description of the 401 is give references 1 and 2.

Terminology. A computer carries out calculations by following a of instructions or *orders* which guide the machine step by step thr the computation. The steps correspond to the very simple opera that computers can perform. The process of breaking down numerical solution of a problem into these steps and the coding o corresponding operations into orders is called *programming* the prob The complete set of instructions, together with any constants use the calculation, is termed the *programme* or the *routine*.

With the 401, instructions and numbers, the latter including da well as constants used in the calculation, are stored in binary form a magnetic disc. The disc or *store* has a capacity of nearly 3 000 or or numbers. Examples of the length of a few programmes are: calculating the logarithm of a number, 51 locations in the store; sine of a number, 30 locations; the square root, 55 locations; randomised block, Latin square, and 3^3 factorial routines referre below occupy 1 000, 1 100, and 1 500 locations respectively.

Orders and numbers are punched on to paper tape, using a printer punch, and are fed into the store by means of a photo-ele reader. An additional form of *input*, that of punched cards, is b incorporated. *Output* of results is by electric typewriter or telepri punch—in the latter case the resulting paper tape can be printed away from the computer. It is intended to employ a tape-to- converter for transferring results obtained from the computer o punched cards.



Elliott-NRDC 401 electronic computer

Types of Problem Discussed. It is convenient to divide our work on the mputer into five main sections:

. Computations previously carried out on desk calculators;

. Problems in which the computations are laborious and time-onsuming and which consequently were neglected when desk nachines alone were available;

. Mathematically intractable problems which can be tackled by nethods involving a huge amount of computation;

. Analysis of survey data;

. Operational routines; these routines are concerned with general roblems of programming and fault-finding.

s was to be expected, the earliest requests to programme work for computer came from our own staff in charge of the desk calculators from external groups, who usually had one tedious computation had to be repeated at frequent intervals. As the benefits of high-d computers have come to be realised so the numbers of requests e grown, and we are now receiving increasing numbers of problems ll descriptions.

he various types of problem will be illustrated by discussing repre-ative examples from each of the sections. (The following members Rothamsted Statistics Department have written programmes men-ed in this paper: F. Yates, J. C. Gower, M. J. R. Healy, S. Lipton, W. Sharp, and D. H. Rees. In the text the initials after the titles rogrammes indicate their authorship.) Clearly a detailed account ld assume a knowledge of programming and of the rather spe-sed order code of the 401. Such detail would be of limited interest has been avoided.

putations Previously Carried out on Desk Calculators

nalysis of the Results of Replicated Experiments. Programmes of this provide analyses of variance in the customary standard form, ther with summaries of the results with their standard errors. Such rammes are necessarily rather elaborate and are only worth ing for designs which are in common use. 'Design' routines at ent available include randomised blocks (M.J.R.H.) of up to 64 ks or 64 treatments with a maximum of 128 plots; Latin squares .) of up to 10×10; the 3^3 factorial design in blocks of 9 with a e replicate (F.Y.); 2^n factorial designs (F.Y.) of up to 128 plots, uding confounding (partial or complete), the estimation of error high-order interactions, fractional replication, factors at 4 or 8 s, and 8×8 quasi-Latin squares. Routines for analysing various s of split-plot designs and 9×9 quasi-Latin squares have also written.

though not presenting any particular theoretical problems, design nes are quite tricky to write. The routine, if it is to be of real use,

must produce a complete analysis taking into account all com requirements of the user. Many minor but irritating difficulties a For example, the numbers being handled can become very large care with *overflow* (exceeding capacity) is required; on the other ha accuracy must be protected, which usually entails at some stage sca the numbers up by a power of two.

Experience has led us to incorporate many checks in the routi In the main these are safeguards against transient machine faults. symbol signifying success or failure of each check is printed. If a ch fails, the machine is instructed either to repeat the calculations lead to that check or, more usually, to print out the discrepancy and carry on. Information as to what has gone wrong is frequently given these discrepancies and the (incorrect) final results. This procedure been found to be better than programming the machine to stop as so as a check fails.

Some of the advantages of using electronic computers for analysi field experiments, as compared with desk machines, are:

1. The results can be obtained in a shorter time. The effect of t has been the elimination of the accumulation of work that used build up during the winter owing to large numbers of experime coming in to be analysed at about the same time.

2. The entire analysis, including the printing in the correct lay can be programmed so that the typed sheets of results produced the computer can be despatched immediately without intermed clerical work.

3. Additional results can be provided. For example, all our de routines calculate the residuals of the observations. Experimen have welcomed having these values available. On desk machines calculation of residuals is impracticable as a routine measure.

4. Programmes can be modified and extended to deal with ad tional requirements, such as missing plots and determination of s of products for covariance work, which are particularly troublesc on desk machines.

This last point is particularly important. As a general rule we h not found it difficult to alter a design routine to take account of so additional complication. A new technique may have to be develop for dealing with the additional work, but design routines are ve flexible and the additions can be programmed in only a fraction of t time that was employed in writing the original routine. The exam of missing observations illustrates this. The standard method for d machines, which necessitates a different iterative formula for ea design, is unsatisfactory for an electronic computer and consequentl new technique[3] was developed: this uses a general iterative proced applicable to all designs and iterates on all the missing values sim taneously.

characteristic of the analysis of experiments is that although the spent by the computer in calculating the results is small, the time to punch out the results is relatively large, e.g. a 6×6 Latin re takes about twelve seconds computing time plus about five ds for reading in data, but the punching out of the analysis takes one minute. Hence the design routine must be constructed so as low the suppression of unwanted parts of the output.

ut of Data. A problem closely associated with the analysis of field iments, but which arises in many other contexts, is the input of Usually an experiment has several sets of data relating to it. For ple, in a grassland experiment the sets of data may be the yields, he percentages of dry matter, crude protein, and phosphate: in a 6 Latin square there would be thirty-six observations in each of our sets. However, in most cases only one or two of the sets of nal data are analysed; functions of the sets are actually required. ne above example neither the yields nor the percentages of dry er are analysed; the chief interest lies in the yield of dry matter, his is obtained as the product of corresponding observations in the sets multiplied by a conversion factor to bring the results into entional units. In other cases the function required may be addi- subtraction, etc., either of corresponding observations of sets or of et and a constant. The design routine requires that each set of undergoing analysis should be in a definite position in the store e computer. The problem, therefore, is to obtain the required ions from the original data and then to place these functions in orrect part of the store.

nless special precautions are taken, a large number of small data- are liable to be scattered about the machine room and an appre- e number of different programme-tapes for forming the functions kely to be required. There is the possibility of losing one of the tapes, of forming the wrong function, of putting the resulting ion in the wrong store locations, and of identifying a set of results the wrong data-tapes.

ur method of dealing with this problem is to have one input ne capable of forming all the commonly required functions. This ed for all the design routines. The original data are punched in for which the method of punching is standardised, all sets for one iment being on one tape with the sets separated by two or three *l orders*. These controls serve as signals to the input routine, indi- g the functions required and the locations to be occupied. This routine, called General Input for Experimental Designs (S.L.) des five services:

It prints out the reference number punched at the head of the periment tape and inputs the original data.

It checks that the correct number of observations are punched in ch set.

3. It forms the maximum, minimum, and mean of each set and p them out. This gives a check on gross punching errors.

4. It forms the functions and prints out their maximum, minin and mean.

5. It places the functions in the store and prints out the loca occupied.

By having one input routine followed immediately by the ana without any intermediate steps, we greatly reduce the chance o operating error. There is one and only one tape for each experim and at the head of the analysis sheet there is printed informa regarding what was analysed. The responsibility for ensuring that correct data are in the right locations and that the correct functions formed is shifted away from the machine operator, who is working w all the noise and bustle inherent in a machine room, to the perso charge of the experiment, who, in the quiet of his own room, dec what control orders are required for the experimental tape. The c trols are kept simple and no knowledge of programming is require construct them.

Fitting of Probit Lines. This topic is characteristic of biological a with a quantal (all or nothing) response. Suppose that succes batches of insects are exposed to varying concentrations of an insecti and that the proportion killed at each concentration is plotted aga the logarithm of the concentration. The resulting curve is sigmoid a transformation to normal equivalent deviates is made and 5 is ad to each value to ensure that all are positive, the resulting values, Y, termed probits. If these are plotted against the logarithms of the c centrations a straight line called the probit line commonly results.

Writing the equation of the line in the form $Y = \bar{y} + b(x - \bar{x})$, problem is to obtain the maximum likelihood estimates of $\bar{y}$, b, an These can be obtained as the result of an iterative process involvi weighted regression calculated at each cycle of the iteration. In instance, the programme (M.J.R.H.) follows closely the standard p cedure used on desk machines. The weights and other quant (working probits) required at each cycle are obtained by a 4-p interpolation into tables held in the store. Up to sixteen concentrati can be handled.

The mode of operating the routine is as follows:

1. Read in the programme and tables of weights and work probits appropriate to the response rate observed in the controls.

2. Read in the data-tape consisting of the logarithm of the conc tration of each dose (x), the number of subjects at each dose, numbers responding, and the initial approximations to the Y's, one each x. It would be possible for the machine to calculate its c initial approximations: this has not been allowed for, in order to courage the biologists to look at their results before sending them for computation.

. On starting the programme the initial Y's are printed. The new proximations obtained by iteration are then computed and printed. s process continues until the convergence is judged to be satis-tory.

. By operating the hand switches the operator causes the machine print the χ^2 value and its degrees of freedom. If the χ^2 value is nificant the operator may decide to allow for this by introducing a erogeneity factor. This is also controlled by the hand switches.

. The routine prints

b	standard error of b
$\bar{x}$ $\bar{y}$	$\bar{y} - b\bar{x}$
ED50	standard error of ED50
	5% fiducial limits of ED50

e ED50 is the estimated dose which will produce the response in 50% he population. If b is not significantly different from zero at the 5% el, it is conceivable that at this level no finite dose will produce a % response; the fiducial limits then include infinity and a string of s is printed.

There are two main reasons for using a 'table look-up' for the weights l working probits rather than computing the probit transformation ues directly. Firstly, owing to the form of the transformation func-n, these values are not easily calculated on a computer. Secondly, nstead of probits a different transformation such as logits or angles required, it is merely necessary to substitute the relevant sets of les; the same routine is valid without alteration. It is possible that alternative procedure to the standard desk machine method could e been developed, but this would have taken considerable time. As re was an urgent need for some programme to carry out probit lysis, it was decided to follow existing techniques.

Recently the procedure has been generalised and a routine for fitting robit plane for the bi-variate case (R.W.S.) is available.

Multiple Regression. This programme (M.J.R.H.) is organised so that possible to examine the linear regressions of any one of a set of up thirty variates on any combination up to a total of nine of the nainder. The initial data read into the computer consist of the ariance matrix of the variates under examination. This covariance trix can be easily obtained from other routines (M.J.R.H. and .G.).

At any stage there are three possibilities for the operator:

. To add a new independent variate to those already used.

. To replace the last independent variate added by another.

. To remove any independent variate from the regression already btained.

the end of each step the machine prints out the residual degrees of

freedom and mean square, and the regression coefficients with t
standard errors.

Heavy Computational Problems

Multivariate Analysis. The following example (M.J.R.H. and S
is presented to exhibit a use of the electronic computer in analys
multivariate data. Detailed results of the computations, together w
a discussion of the anthropological aspects of the work, are be
published elsewhere.

Given a set of measurements on a tooth of a fossil belonging to
unknown species, the problem is to determine the similarities a
differences between this tooth and the corresponding teeth of mod
man and of the great apes. It must be emphasised that no attemp
made to allocate the tooth to any particular species. Indeed, it is cl
from other considerations that the fossils are almost certainly neit
human beings nor members of the existing species of great ape, and
this reason the ordinary discriminant function in its original aspec
unsatisfactory.

It was decided to compare the fossil teeth with those of three 'ra
of living man (Modern British, West African native, Austral
aboriginal) and of both sexes of the chimpanzee, gorilla, and ora
outang, making nine populations in all. A sample was taken from e
population and a number, N, of measurements were made on the sa
type of tooth of every member of each sample. We thus had n
multivariate populations with N variates in each population. In
eight different types of teeth were sampled in this manner, leadin
eight separate analyses and comparisons. N varied from two for
permanent upper second premolar to eight for the permanent lo
canine. The sample size varied for different populations and for
different teeth, ranging between 20 and 190. Altogether the d
comprised about 50 000 measurements.

The technique employed for the comparison is a slight extensio
one described by Rao.[4] If we confine ourselves to one type of to
the method allows us to construct a set of N orthogonal linear functi
of the measurements which are analogous to the ordinary discrimin
function, and which have the property that each taken in turn spa
out the means of the individual populations. Using two of these fu
tions, the results for the populations can be plotted on a two-dimensio
picture. On this we can plot the population means and surround th
by contours which are expected to contain (say) 90% of the valu
belonging to the nine different groups. We can then plot the fossils
the same diagram and so obtain an overall picture of the situati
The method provides criteria for deciding whether two linear functi
are sufficient for specifying the populations without material loss
information and also for deciding whether the points corresponding
the fossils do in fact lie in or near to the plane.

The first stage of the computation involved calculating the mea

ances, covariances, and correlations for each population, and a ine was written for this purpose. The technique used in the comsons requires a common dispersion matrix for the populations, but ection of these results indicated that this was not the case with our . However, the correlation matrices were reasonably constant, and as found that a logarithmic transformation gave dispersion matrices ch were approximately equal. Consequently the variances and ariances had to be recomputed. This presented no difficulty, since ort addition to the original programme enabled the logarithmic sformation to be made as each measurement was read in.

he next stage was the solution of the equation $(\mathbf{B} - \lambda\mathbf{W})\mathbf{x} = 0$, re $\mathbf{W}$ is the common within-population dispersion matrix and $\mathbf{B}$ the veen-population dispersion matrix, the latter being computed from means of the measurements in the different populations. The ropriate linear functions are provided by the solutions ($\mathbf{x}$) of this ation, and if all but two of the latent roots (i.e. λ's) are small in nitude, then the populations can be represented in the two-dimenal space defined by the two corresponding linear functions[4] (Rao rs to the linear functions as canonical variates). The method of solving equation was first to apply the Cholesky reduction to express $\mathbf{W}$ in form $\mathbf{U}'\mathbf{U}$, where $\mathbf{U}$ is upper triangular, so that the equation could written as $(\mathbf{U}'^{-1}\,\mathbf{B}\,\mathbf{U}^{-1} - \lambda\mathbf{I})\mathbf{x} = 0$, where $\mathbf{U}'^{-1}\,\mathbf{B}\,\mathbf{U}^{-1}$ is symrical and $\mathbf{I}$ is a unit matrix. The rotation method[5] was applied nd the latent vectors of $\mathbf{U}'^{-1}\,\mathbf{B}\,\mathbf{U}^{-1}$. Finally multiplication of the nt vectors by $\mathbf{U}^{-1}$, itself given by the Cholesky reduction, provided required linear functions. The programme for this multiplication also used to compute the means and variances of all the correding canonical variates.

he criterion for deciding whether a fossil tooth could be depicted ne same plane as the populations was obtained by calculating not the values for the two selected linear functions but also the values all the remaining canonical variates. These latter values were pared with the corresponding population values, and if there was ppreciable difference then the points representing the fossils were n to be in or near the plane.

will be seen that the amount of computation necessitated was very e, namely:

Computation of means, covariances, and correlations of original ata.

Repetition of (1), using logarithmic transformed data.

Computation of the common dispersion matrix and between-opulation dispersion matrix.

Cholesky reduction of $(\mathbf{B} - \lambda\mathbf{W})$ to obtain $\mathbf{U}'^{-1}\,\mathbf{B}\,\mathbf{U}^{-1}$ and $\mathbf{U}^{-1}$.

Determination of latent roots and vectors of $\mathbf{U}'^{-1}\,\mathbf{B}\,\mathbf{U}^{-1}$ and ultiplication by $\mathbf{U}^{-1}$.

6. Computation of means and variances of canonical variate: populations.

7. Computation of values of canonical variates for fossils.

This work was repeated for each of the eight types of teeth.

The time for these operations is difficult to specify exactly, as it va: from tooth to tooth. The computation of the means, variances, etc., each of the eight teeth took about 90 minutes, of which the lar portion was spent in reading-in data. The between-population disp sion matrix was computed in about one minute. The Cholesky red tion and determination of latent vectors took about fifteen minu However, these operations were carried out in stages. For each st except (1) and (2) previous results had to be punched on to ta Having obtained the variance-covariance matrix, the remaining sta were concluded in about two hours.

Determination of Sampling Errors of Numerical Sequences. The program was developed to deal with an example from the nutritional field bu applicable for all such sequences. The problem was as follows. Gi an individual's dietary intake each week over a period of sixty-th weeks, to estimate the sampling errors which are to be expected sampling further material of the same kind, e.g. taking observati over two consecutive weeks in each successive monthly period. work has been reported elsewhere[5] and will not be elaborated her

The important point is that a method has been developed for tackl such problems which both copes with the statistical difficulties an also suitable for an electronic computer.

Mathematically Intractable Problems

The Combination-of-Probabilities Test of Significance. As in most pr lems of this type, the computing scheme involved is very complica Here, only the problem itself and a brief outline of the method adop for the investigation are given. It is intended to publish a more deta account later.

Let a set of N experiments be performed, each providing an estim $m_1, m_2, \ldots, m_N$ of some parameter μ. If the estimates are all norm distributed with variances $1/W_1, 1/W_2, \ldots, 1/W_N$, then the f efficient overall test of significance is provided by the quantity

$$\left\{\sum_{n=1}^{N}(W_n m_n)\right\} \Big/ \left\{\sum_{n=1}^{N}(W_n)\right\}^{\frac{1}{2}}$$

and this test will result in an exact probability P_L. Alternatively, exact level of significance can be calculated for each of the N estim and the resulting probabilities can be combined to give an ove probability P_C. The latter method, the so-called combination probabilities test of significance, depends on the result that if P_r is exact probability arising from the r'th experiment, then

$$-\sum_{r=1}^{N}(2 \log_e P_r)$$

stributed as χ^2 with $2N$ degrees of freedom. In any particular case nay differ considerably from P_L, and one of the purposes of the k undertaken (F.Y. and D.H.R.) was to compare the relative iency of the two tests.[6]

he technique used in the investigation consisted of fixing N, μ and set of weights W and, by a sampling procedure using random nal deviates, generating a frequency distribution for the combina-of-probabilities test. This was repeated using all permissible pings of four values of N, three values of μ, and five sets of W. Since corresponding distributions of the fully efficient test are known, a parison of the two sets of distributions determines the relative iency of the two tests. The main points of the computing scheme be summarised as follows:

The machine is supplied with a set of 256 random normal viates, four sets of χ^2-tables (one for each value of N), and a table the logarithm of the normal integral.

The machine takes each combination of N, μ, and set of weights be used in turn, and performs the following computations.

For each deviate x the machine computes $y = \log P_N(\mu\sqrt{W} + x)$, here $P_N(z)$ is the normal integral corresponding to the deviate z. four-point interpolation is used for this.

The y's are divided into groups of N and each group is summed. he method of division depends on N and is based on a set of ortho-nal 16×16 Latin squares.

Using the χ^2-table, the machine tests where each of the totals lies the frequency distribution and adds a 1 to the corresponding cell. he total in each cell together with the cumulative sum from one d are finally printed out.

The signs of x are reversed and computations 3 to 5 repeated.

Analysis of Survey Data

a great deal of the survey data we wish to analyse exists at present unched cards, it has been impracticable without having a punched input unit to apply the computer to the analysis of such data. necessary equipment is now being installed so that work with the puter on survey data will begin in the near future.

though some of our present ideas concerning the role of the com-r in this type of work may have to be modified as experience is ed, there appear to be two obvious ways in which the computer assist in obtaining better analyses than are possible at present. ly, the preliminary examination of the data for gross errors, nsistencies, etc., will be more thorough than is now customary or ble. This will be due to the great speed with which computers can

systematically search through data, testing for a variety of diffe contingencies. An example of this sort of work is provided by our of the machine for checking the orders of a skeleton programme certain types of programming error.[2] Secondly, the computer perform arithmetical operations on the data very quickly. This permit the computation of many survey parameters which are present not calculated, as the labour involved is too great and ta too long.

Whether it is desirable to carry out the whole of the analysis on t computer is open to question. Operations such as tabulation can very effectively performed on punch card machines, and a bet procedure may be to punch on to cards intermediate results obtain from the computer and to use punch card machines for part of t work. Whatever experience shows to be the best procedure, the co puter should play an important part in the better analysis of surv data.

Operational Routines

Operational routines are not themselves of a statistical nature, h they are of great service to programming all types of problems and have found it advisable to devote considerable attention to them.

One of our aims has been to make the computer do as much possible of the more routine parts of the work of programming, and experience is that a good deal of the programming can be carried by the computer. A programme has been developed, called the Au matic Programming Routine,[2] which allows programmes for the 4 to be written in a simple code. The routine recodes this skele programme into the orthodox form. Although it took a considera time to write, the routine has already proved its value, both in savi programmers' labour and in producing programmes which contain fewer errors than those in which all the programming is done by ha

Other programmes in this group are:

1. A general output routine (S.L.) for printing numbers as fractio integers, in mixed number form, or in scaling form (the latter pri with the decimal point in the correct position a number scaled up a power of ten). This routine is used for the printing in all the desi routines.

2. *Floating-binary* routines (S.L. and D.H.R.) which avoid m overflow problems.

3. A routine for printing the instructions of a programme in logi order, i.e. for the most part in the sequences in which they are obey (S.L.). In the 401, instructions are not stored in the disc in this mann and a complex routine is needed to print them in this form. Duri the printing the logical faults of the programme are exhibited and t routine has proved valuable in detecting programming errors.

REFERENCES

ON, S. (1955). 'A note on the electronic computer at Rothamsted', *Math. Tab., Wash.*, **9**, 69.

ES, F. and LIPTON, S. (1957). 'An automatic programming routine for the Elliot 401', *J. Assoc. Computing Machinery*, **4**, 151.

LY, M. J. R. and WESTMACOTT, M. H. (1956). 'Missing values in experiments analysed on automatic computers', *Appl. Statist.*, **5**, 203.

, C. R. (1952). *Advanced Statistical Methods in Biometric Research*. Wiley, New York.

SEHOLDER, A. S. (1953). *Principles of Numerical Analysis*. McGraw-Hill, New York.

RCH, B. M. and LIPTON, S. (1956). 'The use of an electronic computer in the estimation of sampling errors in a nutritional survey', *Brit. J. Nutrit.*, **10**, 27.

ES, F. (1955). 'A note on the application of the combination of probabilities test to a set of 2×2 tables', *Biometrika*, **42**, 404.

'THREE-GROUP' REGRESSION ANALYSIS*

PART I. SIMPLE REGRESSION ANALYSIS

WENDY M. GIBSON and GEOFFREY H. JOWETT

University of Sheffield

Investigators who use regression analysis sometimes comp that the standard least-squares method is laborious, especi when no calculating machine is available. In this article Gibson and Dr Jowett develop a method originally describe Professor Bartlett, and show that much information ca quite simply obtained.

In a paper more concerned with the fitting of a linear functio relationship to two variates both subject to error than with regressi Professor M. S. Bartlett[1] described a simple method of estimating regression line of y on x. This consisted of dividing the points on scatter diagram into three equal groups by lines parallel to the y-a and constructing a line through the usual mean point $O(\bar{x}, \bar{y})$ of the observations parallel to the line joining the mean point (A) of group of observations with low values of x to the mean point (B the group with high values of x. The method was an improvemen that given by Professor A. Wald,[2] who had divided the scatter diagr into only two groups. For evenly spaced values of x, Bartlett sho that his method had the surprisingly great efficiency of 89 per c when compared with the least-squares method. In this paper, an another to follow, Bartlett's method is developed and extended multiple regression with two independent variables, and easy meth are given for estimating the large-sample standard errors of the co cients. These techniques involve much less arithmetic than ordin least-squares regression and are quick and accurate, the time ta using no calculating machine being about the same as the time ta by the least-squares method using a good semi-automatic electri machine. There are also useful by-products in the form of grap which help to illuminate the relationships.

We feel that an exposition of this technique will be useful to th classes of people: (*a*) field workers, as a substitute for the conventio method when they have no automatic calculator available; (*b*) ind trial workers with non-mathematical training who feel ill at ease w the mathematics of the usual method and who may find this met more 'common-sense' in character; and (*c*) teachers and students, w

* Read in part to the Sheffield Group of the Industrial Applications Section of the R Statistical Society, October 1956.

find in the theory of the easy method not only an illuminating duction to ordinary regression theory, but also an illustration of al important statistical ideas and techniques, such as efficiency, inearity of regression, homoscedasticity (constancy of dispersion), stness of tests, combination of standard errors, interpretation of ssion planes, and so forth.

e procedure with one independent variate will be illustrated with ence to a medical example, for which the data are given in Table I a scatter diagram is given in Fig. 1.

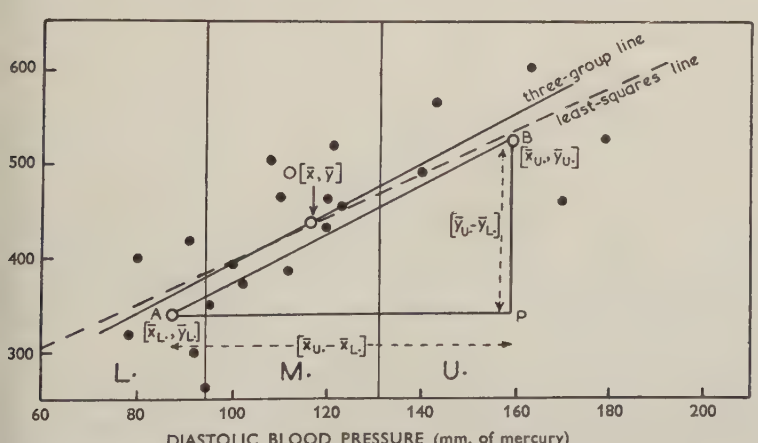
TABLE I

l data: diastolic blood pressure in mm. of mercury (x) and heart weight in grams (y) in men dying of cerebral haemorrhage

x	y	Group	x	y	Group
121	521	M·	143	567	U·
120	465	M·	94	264	L·
95	352	M·	108	503	M·
123	455	M·	78	320	L·
140	490	U·	170	458	U·
112	388	M·	80	399	L·
92	301	L·	179	524	U·
100	395	M·	110	465	M·
102	375	M·	120	432	M·
91	418	L·	163	602	U·

g the Line

e points on the scatter diagram are first divided, by means of parallel to the y-axis, into three groups, L·, M·, U·. In this nota- L·, for example, signifies the group of points containing low values rrespective of the values of y (·L will be used to signify low values irrespective of the values of x). M, U signify middle, upper ctively. The groups are chosen so as to contain respectively $n_{L\cdot}$, $_{U\cdot}$ points, where the ratios $n_{L\cdot} : n_{M\cdot} : n_{U\cdot}$ may be chosen arbitrarily,



Scatter diagram of data in Table I, showing construction of three-group regression line.

but are usually best taken as approximately 1 : 2 : 1. The choic these ratios is not very critical, and will be discussed more fully p. 119.

TABLE II

Computation of b_{yx} from medical data

Group	n	Totals		Means		$b_{yx} = \frac{\bar{y}_{U\cdot} - \bar{y}}{\bar{x}_{U\cdot} - \bar{x}}$
		x	y	x	y	
L·	5	435	1 702	87·00	340·40	187·80
M·	10	1 111	4 351	(not required)		72·00
U·	5	795	2 641	159·00	528·20	= 2·608
Total	20	2 341	8 694	117·05	434·70	—

The means $(\bar{x}_{L\cdot}, \bar{y}_{L\cdot})$, $(\bar{x}_{U\cdot}, \bar{y}_{U\cdot})$ for the outer groups are then co puted as shown in Table II and plotted as points A, B on the scat diagram as shown in Fig. 1. The regression line is then drawn throu the overall mean point $O(\bar{x}, \bar{y})$ in a direction parallel to AB; its sl is thus given by

$$b_{yx} = (\bar{y}_{U\cdot} - \bar{y}_{L\cdot})/(\bar{x}_{U\cdot} - \bar{x}_{L\cdot}) \quad \ldots$$

and its equation by $y - \bar{y} = b_{yx}(x - \bar{x})$. For the medical data thi $y - 434{\cdot}70 = 2{\cdot}608(x - 117{\cdot}05)$, and is shown as a full line in Fig The least-squares regression line has equation

$$y - 434{\cdot}70 = 2{\cdot}263(x - 117{\cdot}05)$$

and is shown for comparison by the broken line in Fig. 1.

In most practical applications only one of the regression line relevant, but if the regression line of x on y is required it is given $(x - \bar{x}) = b_{xy}(y - \bar{y})$, where $b_{xy} = (\bar{x}_{\cdot U} - \bar{x}_{\cdot L})/(\bar{y}_{\cdot U} - \bar{y}_{\cdot L})$, the syml ·L, ·M, ·U denoting the groups obtained by drawing lines parall the x-axis.

Measuring Residual Scatter

In practical applications the standard error of the regression coe cient b_{yx} is often required, and for this an estimate of the stand deviation of the y-values about the regression line is needed. It wo be permissible to calculate the actual root-mean-square of the dev tions as in least-squares regression, but two alternative methods, wh are less laborious and almost as good, are described below. These m be used when residual scatter may be assumed to be normally d tributed.

(i) *An Estimate of σ_0 based on Range.* Quasi-random samples of dev tions from the regression line may be obtained by dividing the diagr into bands parallel to the y-axis, each containing about the sa number of points (less than twelve, since the use of range is unsa

ry except in small samples), as shown in Fig. 2. The range of ations from the regression line in each band is then found by ing lines parallel to the regression line through the extreme points ither side of it in the band and measuring the vertical distance een them, as in Fig. 2. Each distance is then divided by the usual

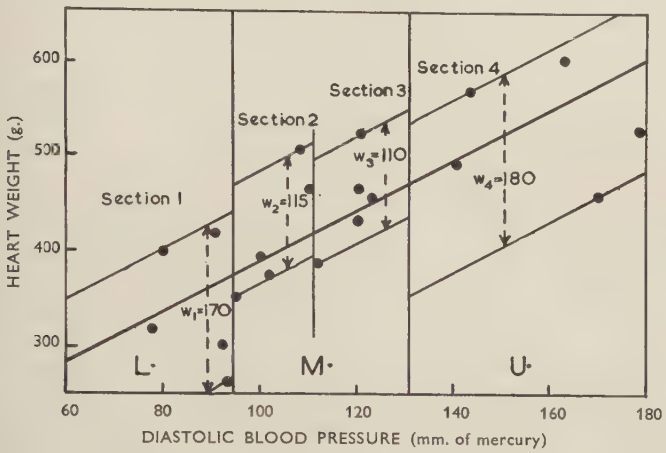


FIG. 2. Measurement of ranges for estimation of σ_0; data from Table I.

r, d, to convert it into an estimate of σ_0, and the estimates obtained the different bands are averaged to give a single estimate as n in Table III. By this method the data of Table I yield an ıate of 61·8 for σ_0.

TABLE III

Estimation of σ_0 from medical data

Section	n	Range w	Factor d	w/d	Est. σ_0
L·) 1	5	170	2·326	73·1	
M·) 2	5	115	2·326	49·4	$\frac{247·2}{4}$
M·) 3	5	110	2·326	47·3	$= 61·80$
U·) 4	5	180	2·326	77·4	
Total	20	—	—	247·2	—

An Estimate of σ_0 from the Mean Absolute Deviation. In large ɔm samples from a normal population the mean deviation is l, on the average, to $\sigma/1·25$. In regression, the absolute deviations the regression line may be read off the diagram (possibly with id of a ruler), averaged, and the result multiplied by 1·25 to give timate of σ_0. Since the deviations are from the observed, not the regression line, an estimate obtained in this way may be expected biased a little on the low side; it is not a simple matter to estimate ias exactly, but there are some grounds for multiplying the factor

1·25 by $\sqrt{[n/(n-2)]}$ (cf. the factor $\sqrt{[n/(n-1)]}$ used in a sim way by Godwin[3] for an ordinary random sample). By this meth the data of Table I yield an estimate of 62 for σ_0.

Use of the range tends to be quicker and easier when the data numerous, but has the disadvantage of being based on arbitrary bar

Estimation of the Large-sample Standard Error of b_{yx}

In linear regression, y_i is usually assumed to be composed of a p $\alpha + \beta x_i$ attributable to regression, plus a residual quantity ξ_i, wh is randomly distributed with zero mean and standard deviation σ_0.

Hence

$$b_{yx} = \frac{\bar{y}_{U\cdot} - \bar{y}_{L\cdot}}{\bar{x}_{U\cdot} - \bar{x}_{L\cdot}} = \frac{(\alpha + \beta\bar{x}_{U\cdot} + \bar{\xi}_{U\cdot}) - (\alpha + \beta\bar{x}_{L\cdot} + \bar{\xi}_{L\cdot})}{\bar{x}_{U\cdot} - \bar{x}_{L\cdot}}$$

$$= \beta + \frac{\bar{\xi}_{U\cdot} - \bar{\xi}_{L\cdot}}{\bar{x}_{U\cdot} - \bar{x}_{L\cdot}}$$

Now β, being the true regression slope, is constant, so for sets of d with the same values of x_i,

$$\text{var } (b_{yx}) = \frac{1}{(\bar{x}_{U\cdot} - \bar{x}_{L\cdot})^2} \text{var } (\bar{\xi}_{U\cdot} - \bar{\xi}_{L\cdot})$$

Now $\bar{\xi}_{U\cdot}$, $\bar{\xi}_{L\cdot}$ are merely means of two virtually random samples deviations from regression, containing respectively $n_{U\cdot}$, $n_{L\cdot}$ membe therefore

$$\text{var } (\bar{\xi}_{U\cdot} - \bar{\xi}_{L\cdot}) = \frac{{\sigma_0}^2}{n_{U\cdot}} + \frac{{\sigma_0}^2}{n_{L\cdot}}$$

so that

$$\text{var } (b_{yx}) = \left(\frac{1}{n_{U\cdot}} + \frac{1}{n_{L\cdot}}\right){\sigma_0}^2/(\bar{x}_{U\cdot} - \bar{x}_{L\cdot})^2 \qquad \ldots$$

This is always somewhat greater than the variance of the coeffici obtained by least squares, an optimum procedure, by an amo depending on circumstances; for x's approximately normally tributed, and the ratios 1 : 2 : 1, it is about 23 per cent greater, lead to a standard error of b_{yx} which is 11 per cent greater. For the d of Table I est. $\sigma(b_{yx})$ is calculated as 0·543 by the three-group meth using range, compared with 0·479 by the least-squares method. T difference is greater than 11 per cent in this case, because of the of different estimates of σ_0, which might equally well have worked the opposite direction. The difference using a common estimate of would have been just under 11 per cent. Whichever method is us the significance of the regression is clearly established.

Allowing for Variation of Dispersion

If the dispersion of y about the regression line is markedly differ in the three groups (a situation termed heteroscedasticity), as sometir

pens, the method of estimating var (b_{yx}) described above may y be modified to take account of this. Separate estimates $\sigma_{(U\cdot)0}$ $\sigma_{(L\cdot)0}$ of residual scatter may be determined for the upper and r group and we then have

$$\text{var}\,(b_{yx}) = \left(\frac{\sigma^2_{(U\cdot)0}}{n_{U\cdot}} + \frac{\sigma^2_{(L\cdot)0}}{n_{L\cdot}}\right) \Big/ (\bar{x}_{U\cdot} - \bar{x}_{L\cdot})^2 \qquad \ldots . (3)$$

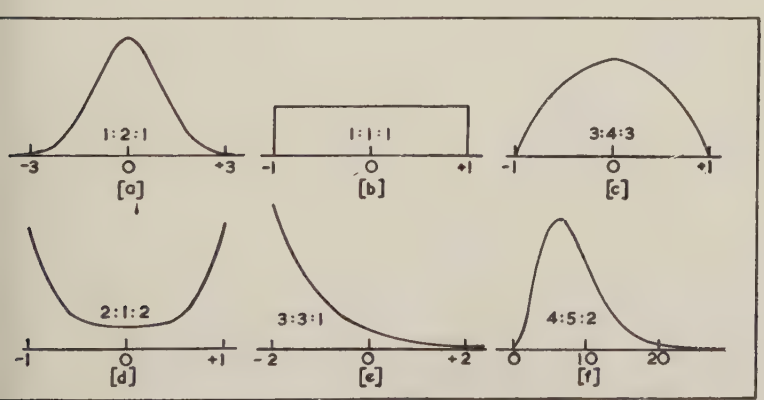
lently formula (2) is robust in Dr G. E. P. Box's sense as regards artures from homoscedasticity; if $n_{U\cdot}$ and $n_{L\cdot}$ are roughly equal, as usually be the case, the numerator in (3) will be roughly

$$\left(\frac{1}{n_{U\cdot}} + \frac{1}{n_{L\cdot}}\right) \tfrac{1}{2}(\sigma^2_{(U\cdot)0} + \sigma^2_{(L\cdot)0})$$

this will usually not be very different from the numerator in (2) σ_0^2 estimated as described in the preceding section.

mum Choice of the Numbers of Points in the Three Groups

ne optimum allocation of points to the three groups varies with type of x-distribution: optimum ratios have been derived for some retical types as shown in Fig. 3. The method of determination the efficiencies, compared with the least-squares method, resulting the choice of various ratios are given in the Appendix and suggest unless non-normality is very marked (i.e. enough to be striking ne scatter diagram itself) little is lost by using the ratios 1 : 2 : 1 standard practice. For serious departures from normality it seems e a general rule that short fat tails need more points, long thin fewer, than the normal quota.



Types of x-distribution, with the optimum ratios in the three groups: (*a*) normal; (*b*) rectangular; (*c*) bell-shaped; (*d*) U-shaped; (*e*) J-shaped; (*f*) skew.

for Linearity of Regression

ne three points used in fitting the regression line may also be used, g lines suggested by Bartlett, to provide a convenient test for rity of regression. If h is the distance in the y direction of the

point O from the line joining A and B, then the comparison of h v twice its standard error provides such a test. For by elemen co-ordinate geometry

$$h = \bar{y} - \bar{y}_{L\cdot}\left(1 - \frac{\bar{x} - \bar{x}_{L\cdot}}{\bar{x}_{U\cdot} - \bar{x}_{L\cdot}}\right) - \bar{y}_{U\cdot}\left(\frac{\bar{x} - \bar{x}_{L\cdot}}{\bar{x}_{U\cdot} - \bar{x}_{L\cdot}}\right)$$

$$= \frac{n_{U\cdot}\bar{y}_{U\cdot} + n_{M\cdot}\bar{y}_{M\cdot} + n_{L\cdot}\bar{y}_{L\cdot}}{n_{U\cdot} + n_{M\cdot} + n_{L\cdot}} - \frac{\bar{x}_{U\cdot} - \bar{x}}{\bar{x}_{U\cdot} - \bar{x}_{L\cdot}}\bar{y}_{L\cdot} - \frac{\bar{x} - \bar{x}_{L\cdot}}{\bar{x}_{U\cdot} - \bar{x}_{L\cdot}}\bar{y}_{U\cdot}$$

$$= \bar{y}_{U\cdot}\left(\frac{n_{U\cdot}}{n} - \frac{\bar{x} - \bar{x}_{L\cdot}}{\bar{x}_{U\cdot} - \bar{x}_{L\cdot}}\right) + \bar{y}_{M\cdot}\left(\frac{n_{M\cdot}}{n}\right) + \bar{y}_{L\cdot}\left(\frac{n_{L\cdot}}{n} - \frac{\bar{x}_{U\cdot} - \bar{x}}{\bar{x}_{U\cdot} - \bar{x}_{L\cdot}}\right) \ldots .$$

and so

$$\sigma_h^2 = \frac{\sigma_0^2}{n_{U\cdot}}\left(\frac{n_{U\cdot}}{n} - \frac{\bar{x} - \bar{x}_{L\cdot}}{\bar{x}_{U\cdot} - \bar{x}_{L\cdot}}\right)^2 + \frac{\sigma_0^2}{n_{M\cdot}}\left(\frac{n_{M\cdot}}{n}\right)^2 + \frac{\sigma_0^2}{n_{L\cdot}}\left(\frac{n_{L\cdot}}{n} - \frac{\bar{x}_{U\cdot} - \bar{x}}{\bar{x}_{U\cdot} - \bar{x}_{L\cdot}}\right)^2 \ldots .$$

For the medical data $h = 16{\cdot}0$ and σ_h is estimated, using (5), as 14 there is therefore no evidence of departure from linear regressi Such a test is, of course, only a check on the collinearity of $(\bar{x}, \bar{y})$, ($\bar{y}_{U\cdot}$), $(\bar{x}_{L\cdot}, \bar{y}_{L\cdot})$; it would tend to detect, say, a 'parabolic' depart from linearity, but not a 'cubic' type of departure such as might le these three means in a straight line.

Procedure when Points fall on Dividing Lines

If a dividing line passes through a number of points, it is best move it to one side or the other so as to secure the smaller value of factor $(1/n_{U\cdot} + 1/n_{L\cdot})/(\bar{x}_{U\cdot} - \bar{x}_{L\cdot})^2$ in (2). The efficiency of the reg sion procedure is not sensitive to slight changes in proportions Appendix), and so it is not usually necessary to follow the more ela rate but theoretically more desirable alternative procedure of count a fraction of every borderline point in both groups and modifying standard error formula by adding the sum of squares of such fracti to the appropriate number $n_{U\cdot}$ or $n_{L\cdot}$ in (2).

APPENDIX

Derivation of Optimum Proportions

The derivation of optimum proportions was carried out for following six distributions of x; they are, of course, unaffected by m and scale. The corresponding frequency curves are shown in Fig.

Normal:	$f(x) \propto e^{-x^2/2}$,	$-\infty < x < +\infty$
Rectangular:	$f(x) \propto \frac{1}{2}$,	$-1 < x < +1$
Bell-shaped:	$f(x) \propto 1 - x^2$,	$-1 < x < +1$
U-shaped:	$f(x) \propto 1/4 + x^4$,	$-1 < x < +1$
J-shaped:	$f(x) \propto e^{-x}$,	$-2 < x < +\infty$
Skew:	$f(x) \propto x^3 e^{-x/2}$,	$0 < x < +\infty$

ıe method was to obtain an expression for var (b_{yx}) in terms of the ssae θ_1, θ_2, of the dividing lines, and then minimise it, either by rentiation, or by an iterative method. The proportions under the ency curves corresponding to these abscissae were then deduced, are given in Table IV. The efficiency of the three-group method

TABLE IV

imum proportions for different distributions of x for use with the three-group method

Distribution	Proportions n_L. : n_M. : n_U.	Approximate Ratios n_L. : n_M. : n_U.
rmal	0·27 : 0·46 : 0·27	1 : 2 : 1
ctangular	0·33 : 0·33 : 0·33	1 : 1 : 1
-shaped	0·31 : 0·38 : 0·31	3 : 4 : 3
haped	0·39 : 0·22 : 0·39	2 : 1 : 2
aped	0·45 : 0·40 : 0·15	3 : 3 : 1
w	0·36 : 0·45 : 0·19	4 : 5 : 2

ve to the least-squares method was then determined for each type stribution by expressing the variance of the least-squares estimate $_x$ as a percentage of that of the optimum three-group estimate. any distribution with frequency density proportional to f(x),

$$\text{var}\,(b_{yx}\,[\text{least squares}]) = \sigma_0^2 \Big/ \int_{-\infty}^{+\infty} (x - \bar{x})^2\, f(x)\,dx$$

$$b_{yx}[\text{three-group}]) = \frac{\sigma_0^2 \left(1 \Big/ \int_{-\infty}^{\theta_1} f(x)\,dx + 1 \Big/ \int_{\theta_2}^{+\infty} f(x)\,dx \right)}{\left[\left(\int_{-\infty}^{\theta_1} x f(x)\,dx \Big/ \int_{-\infty}^{\theta_1} f(x)\,dx \right) - \left(\int_{\theta_2}^{+\infty} x f(x)\,dx \Big/ \int_{\theta_2}^{+\infty} f(x)\,dx \right) \right]^2}$$

e θ_1, θ_2 are calculated as described above. It would be more enient, even at the expense of a slight loss of efficiency, to have ne set of ratios for use in all cases; accordingly, for each distribu- the efficiency of the optimum normal and rectangular ratios ve to the optimum ratios was determined and is given in Table V.

TABLE V

Relative efficiency of three-group method with different ratios

istribution of x	Optimum Three-group Relative to Least Squares	1 : 2 : 1 Three-group Relative to Optimum Three-group	1 : 1 : 1 Three-group Relative to Optimum Three-group
mal	81	100	98
tangular	89	97	100
-shaped	86	99	99
haped	93	87	97
aped	79	81	78
v	80	93	91

It will be seen that the optimum three-group method is surpris efficient, and that the ratios used are not very critical except in of extreme skewness. If all these distributions occurred with equa quency in practice, there would be little to choose between the r 1 : 2 : 1 and the equal numbers advocated by Professor Bartlett; since the normal distribution is the most common type, excep evenly spaced data which are easy to analyse by least squares, it s best to advocate general use of the ratios 1 : 2 : 1.

REFERENCES

[1] Bartlett, M. S. (1949). 'Fitting a straight line when both variables are su to error', *Biometrics*, **5**, 207.

[2] Wald, A. (1940). 'The fitting of straight lines if both variables are subje error', *Ann. Math. Statist.*, **11**, 284.

[3] Godwin, H. J. (1945). 'On the distribution of the estimate of mean devia obtained from samples from a normal population', *Biometrika*, **33**, 254

DISCRIMINANT FUNCTIONS OF SOCIO-ECONOMIC CLASS*

KATHLEEN GALES
Research Techniques Unit, London School of Economics

The assessment of 'social class' is often carried out subjectively, and there is a danger that assessors may differ in their judgements. In this article Miss Gales discusses the use of discriminatory analysis in the assignment of individuals to one of three socio-economic classes, and describes the results of a preliminary investigation of the problem.

any sample surveys interviewers are asked to assess the 'social of respondents. In some organisations they are asked to base assessment primarily on objective criteria such as income and pation, whereas by other organisations a judgement of the 'social of individuals is required. It is known that different social groups different patterns of expenditure, and breakdowns by social class herefore of considerable interest, however nebulous the concept be.

ke most procedures that have a loophole for subjective bias, the fication by interviewers is not entirely satisfactory. It is natural, fore, to consider the possibility of constructing an objective index cio-economic status, based on those observable factors, such as pation, income, education, house-ownership, and so forth, that are ly assumed to be correlated with it. This paper describes the s of a preliminary investigation of this problem, in which the ique of discriminatory analysis was applied to data concerning the cteristics of persons placed by interviewers in three 'social classes'. bject was to discover, for a given set of informant characteristics inear discriminant functions, that level of agreement which could btained between the interviewer classification and the 'best' tive classification.

Theory of Discriminatory Analysis

any application of discriminatory analysis to data concerning the s of p variables taken by members of two or more populations, n functions are constructed. By the use of these functions any idual whose p variate-values are known, but whose parent popula- is unknown, may be assigned to one of the populations, the

uch of the work on this project was carried out by two former members of the Research ques Unit, Mr R. I. Cole and Mrs Erica Royston.

coefficients of the variables in the functions being chosen to maxi the probability of correct assignment. Thus, if we know the valu variables x_1, x_2, x_3 taken by members of populations P_1 and P_2 we use a linear function of the variables:

$$F = k_0 + k_1x_1 + k_2x_2 + k_3x_3$$

to assign individuals of unknown population to P_1 or P_2, choosing k's in such a way that the probability of correct assignment is gre for this set of coefficients than for any other set. The basic problem t consists in the determination of that set of coefficients which, for a gi set of variate values, maximises the probability of correct assignm The greater the extent to which the populations overlap, in the se that certain members could belong to two or more of the populatio the lower will be the probability of correct assignment by even 'best' of functions.

The first published application of discriminatory analysis appea in 1935[1] and concerned the estimation of the archaeological period which a skull of unknown age belonged. The technique has since t repeatedly been applied in anthropology. In the field of taxonor also, ample scope is provided for its use, while in the social scie there have been several applications, including a study of the cha teristics of successful salesmen (in 1938[2]) and of Borstal boys (in 195

If an individual of unknown parent population but known var values $(x_1, x_2, \ldots, x_p)$ is to be assigned to one of three populati P_1, P_2, P_3 which exist in proportions π_1, π_2, and π_3, then, for probability of wrong assignment to be minimised, the following r of procedure must be adopted:

$$\text{assign to } P_1 \text{ if } \pi_1 f_1 \geqslant \pi_2 f_2,\ \pi_1 f_1 \geqslant \pi_3 f_3;$$

$$\text{assign to } P_2 \text{ if } \pi_2 f_2 \geqslant \pi_3 f_3,\ \pi_2 f_2 \geqslant \pi_1 f_1;$$

$$\text{assign to } P_3 \text{ if } \pi_3 f_3 \geqslant \pi_1 f_1,\ \pi_3 f_3 \geqslant \pi_2 f_2$$

where f_1, f_2, f_3 are the probability density functions of the varia $x_1, x_2, \ldots, x_p$ in populations P_1, P_2, P_3.

In the particular case when each population is multivariate normal in the p variates, each having the same dispersion matrix these relations reduce to a fairly simple form. We assign an individ to P_1, P_2, or P_3 according to which of the functions L_1, L_2, or L_3 ta the maximum value when in it are substituted the variate values of particular individual, where

$$L_1 = \sum_i (\sum_j \alpha^{ij}\mu_{1j})x_i - \tfrac{1}{2}\Sigma\alpha^{ij}\mu_{1i}\mu_{1j} + \log \pi_1$$

$$L_2 = \sum_i (\sum_j \alpha^{ij}\mu_{2j})x_i - \tfrac{1}{2}\Sigma\alpha^{ij}\mu_{2i}\mu_{2j} + \log \pi_2$$

$$L_3 = \sum_i (\sum_j \alpha^{ij}\mu_{3j})x_i - \tfrac{1}{2}\Sigma\alpha^{ij}\mu_{3i}\mu_{3j} + \log \pi_3$$

μ_{1j}, μ_{2j}, μ_{3j} $(j = 1 \text{ to } p)$ are the mean values of the variates in P_1,

P_3, and α^{ij} is the inverse of α_{ij}. L_1, L_2, and L_3 are linear functions , x_2, . . . x_p, the constants being functions of the parameters of parent populations P_1, P_2, P_3. From a sample of members from hree populations, these parameters can be estimated and the linear ions constructed.

r example, if values of two characteristics x_1 and x_2 are available nembers of three populations P_1, P_2, and P_3, on evaluation of the icients we might arrive at functions

$$L_1 = 2{\cdot}5x_1 + 2{\cdot}3x_2 + 0{\cdot}4$$

$$L_2 = 2{\cdot}3x_1 + 4{\cdot}2x_2 - 1{\cdot}3$$

$$L_3 = 1{\cdot}8x_1 + 5{\cdot}1x_2 + 3{\cdot}2$$

n if an individual of unknown population has variate values 1, 3 it be assigned to P_3, since the values taken by L_1, L_2, L_3 for $x_1 = 1$, 3 are 9·8, 13·6, and 20·3 respectively.

cation to Social Class

applying discriminatory analysis to social class two difficulties are ediately apparent: the first arises from the impossibility of obtain- ny incontrovertible data on the social class of a sample group of viduals, and the second from the non-measurable nature of the acteristics correlated with class. A panel of experts might perhaps been used to classify respondents, but we did not wish, at this , to undertake a special survey, and no such data were available. as therefore necessary to use as the basic data from which to nate parent parameters a sample of individuals classified by inter- ers. In constructing discriminants from such data, any check on validity' of our final classification was sacrificed.

w of the factors correlated with socio-economic class are measur- continuous normal variates, such as those occurring in anthropo- y and taxonomy. The procedure here followed in order to impose tem of measurement on the data was, in regard to each of the rved factors such as occupation, education, etc., to rank the ble categories (grouped where necessary) and assign to each idual the rank score corresponding to the category into which his bute value fell. Thus, for a simple dichotomy such as house- rship, each individual scored 0 or 1, while for education a fivefold fication was used and a score of −2, −1, 0, 1, or 2 could be ned.

e assumption that each variate was continuous and normally buted was obviously unjustified, and the application of discrimi- analysis to these data can be defended only on the heuristic nd that it is difficult to think of anything better; its ultimate cation in any case must depend on whether or not it is found to use.

We decided to construct discriminant functions from two sets of d the first coming from a quota sample and the second from a rand sample. It was expected that there would be a closer agreem between the objective index and the interviewer judgements for random sample.

Quota Sample Data

Information concerning the characteristics of interviewer-classifi respondents forming part of a quota sample was taken from a rando sub-sample of some of the questionnaires completed in a previo survey undertaken by the Research Techniques Unit.[4] In this surv thirty-six interviewers from three organisations had been employed an informants had been placed in one of three social classes. Data c about 140 males and 180 females from each of these classes were use

There were nine attributes for the males, upon which informatio had been obtained and which were expected to be correlated wi class. These were income; education; occupation; ownership of hous television, telephone, and car; entry in football pools; and newspap readership. Scores were assigned to each respondent, corresponding his nine attributes. Occupation was scored according to the Registra General's five socio-economic groups, education according to five ag of completion, and income according to four ranges. The ownersh attributes were scored 1 or 0 according as the various goods were were not possessed. Similarly, a score of 1 was given to individu betting on football pools and to those reading various 'high-grad newspapers. The same procedure was followed for the females, exce that here only seven attributes were scored, income and occupati being omitted.

Correlation coefficients were calculated between each of the scor attributes and social class (1, 2, or 3) as assessed by interviewe Although the theoretical justification for this procedure is again ve doubtful, the results (shown in Table I) were rather as might have be expected.

TABLE I

Correlation coefficients between social class and various attributes

	Occupation	Income	Education	Ownership of: Television	House	Telephone	Car	Pools Betting	Newspaper Readership
Males	0·701	0·732	0·572	0·162	0·439	0·537	0·472	0·233	0·491
Females			0·596	0·116	0·495	0·517	0·407	0·134	0·452

It was not surprising to find that the correlation of class with occupa tion, income and education gave the highest coefficients, since these a

attributes in terms of which social class is usually defined. The ership of television showed a very low correlation with class and put to no further use in the analysis.

order to construct the linear discriminant functions L_1, L_2, L_3 for males, the mean value μ_{1i}, μ_{2i}, μ_{3i} ($i = 1$ to 8) of each attribute ch class was estimated from its value in the sample. The dispersion rix α_{ij} for the eight attributes in the entire population was also nated from its value in the sample. The eight by eight matrix was inverted by the Doolittle method[5] to obtain α^{ij}. From the product $[\mu_{1i}]$ the coefficients $(X_1, X_2, \ldots, X_8)$ of $(x_1, x_2, \ldots, x_8)$ in L_1 then obtained. The coefficients in L_2 and L_3 were similarly ulated. The constant terms were then calculated, the π's being ted, since the discriminants were to be used only for classifying viduals belonging to the sub-sample, in which the proportions of different classes were roughly equal. A similar procedure was wed for the sub-sample of female respondents, the discriminant tions in this case being based upon six attributes.

he interviewer-classified respondents were then reclassified by the iminant functions in order to discover the level of agreement veen the two methods of classification. The results obtained are vn in Table II.

TABLE II

Comparison of classifications by interviewers and discriminants

Males

Placed by Interviewers in Social Class	Placed by Discriminant in Social Class			Total
	I	II	III	
I	116	20	8	144
II	33	48	60	141
III	14	13	108	135
Total	163	81	176	420

Females

Placed by Interviewers in Social Class	Placed by Discriminant in Social Class			Total
	I	II	III	
I	155	12	21	188
II	79	22	83	184
III	11	4	162	177
Total	245	38	266	549

Since neither classification can be tested against one whic accepted as completely valid, comparisons can merely be made tween what the interviewers thought and what the discrimin suggest. There were obviously large divergences. Tables III and show the 'misclassification' in each case, on the assumption that other classification is correct.

TABLE III

Misclassification by discriminants—assuming interviewers correct

Social Class	% Males Misclassified		% Females Misclassified	
	Actual	Theoretical	Actual	Theoretical
I	19	20	18	31
II	77	20	88	64
III	20	0·3	8	31

TABLE IV

Misclassification by interviewers—assuming discriminants correct

Social Class	% Males Misclassified	% Females Misclassified
I	29	37
II	41	42
III	39	39

The 'theoretical' probability of misclassification depends on extent to which there is similarity between persons placed by in viewers in different social classes, and was calculated from the according to a method of Rao,[6] again assuming multivariate nor populations and continuous variables. A high theoretical percen of misclassification for one class indicates a great degree of simila in respect to the characteristics considered, between that class and adjacent class (or classes).

The divergence between discriminants and interviewers in the cla fication of those individuals placed by interviewers in the second soc class is very striking and is, in the case of males, nearly four times expected value. It might also be noted that the expected errors in t classification of females are distinctly greater than those for males. T suggests the greater similarity of the classes of females in respect their six attributes than of the males in respect of their eight.

From the scores of an individual on each of the variables a total sc can be constructed by simple addition. This simple score can also used to allocate to social classes. A test of the method was found to g approximately the same amount of misclassification as given by discriminants. A classification of males merely according to occupati

ng the Registrar-General's five classes), ignoring all the other ables, gave only a slightly greater deviation from the interviewer ification than did the discriminant method. Actual results ob- ed are shown in Table V.

TABLE V

Levels of agreement between different classificatory methods

Agreement (males) between:	%
Scores and interviewers	65
Discriminants and interviewers ..	65
Occupational classification and interviewers	58
Occupational classification and discriminants	54
Occupational classification and scores	56
Discriminants and scores	70

dom Sample Data

ne random sample data used by us had been obtained by the sh Market Research Bureau in their Derby Survey.[7] In this survey nty-five interviewers had been employed and had placed respon- s in three social classes. Information concerning about forty-six es and forty females from each class was used in the construction scriminant functions. The size of the smallest class was used as a h determinant of the size of the samples from the other classes as as desirable that estimates of population parameters from sampling should be of equal degrees of accuracy.

ne attributes likely to be correlated with social class, upon which rmation was available, were income, education, occupation, self- sed social class, possession of a refrigerator, readership A, reader- B, and voting behaviour. In this case the occupation and income e head of the household were recorded also for females.

ccupation was scored according to five groups, income for six ps, and education for five ages of school leaving. Self-assessed l class was scored 1 to 4 for the groups recorded as Working, er-Middle, Upper-Middle, and Upper respectively. Possession refrigerator gained a score of 1. Readerships A and B were a ure of the type of magazines and newspapers read by the respon- A score of 1 in A indicated that at least one of a certain number rly 'high-grade' publications was stated to have been read in the day, week or month, while a score of 1 in B indicated that none rtain 'low-grade' publications were read. Information was avail- for a very wide range of magazines and newspapers and only those d to be most highly correlated with social class were actually used, rent selections being made for men and women. For the final acteristic, voting behaviour, a score of 1 was given to a conservative and 0 was given to all others.

Correlation coefficients were calculated between each of the ei attributes and the interviewer-assessed 'social classes'. The results shown in Table VI.

TABLE VI

Correlation coefficients between social class and various attributes

	Occupation	Income	Education	Self-assessed Class	Refrigerator	Reading A	Reading B	Voting
Males	0·668	0·744	0·559	0·654	0·207	0·370	0·163	0·528
Females	0·618	0·652	0·598	0·464	0·097	0·467	0·167	0·290

As in the quota sample, the characteristic found to be most high correlated with interviewer classification is seen to be income, wh occupation comes second in order. For males, the correlation of cla with self-assessed class and voting behaviour is of interest.

Discriminant functions based on eight attributes were obtain separately for males and females. The original interviewer-classifi informants were reclassified by these functions, and interviewer a discriminant classifications were compared, as in the case of the quo sample data. The results are shown, together with the calculated the retical errors, in Tables VII–IX.

TABLE VII

Comparison of classifications by interviewers and discriminants

Males

Placed by Interviewers in Social Class	Placed by Discriminant in Social Class			Total
	I	II	III	
I	40	6	0	46
II	13	23	10	46
III	2	7	37	46
Total	55	36	47	138

Females

Place by Interviewers in Social Class	Placed by Discriminant in Social Class			Total
	I	II	III	
I	28	5	0	33
II	9	21	10	40
III	0	11	29	40
Total	37	37	39	113

TABLE VIII

Misclassification by discriminants—assuming interviewers correct

Social Class	% Males Misclassified		% Females Misclassified	
	Actual	Theoretical	Actual	Theoretical
I	13	29	15	28
II	50	46	48	46
III	20	29	27	21

TABLE IX

Misclassification by interviewers—assuming discriminants correct

Social Class	% Males Misclassified	% Females Misclassified
I	39	25
II	43	43
III	29	26

he overall disagreement between interviewers and discriminants ery similar to that in the quota sample case, but is spread more ly over the entire sample of informants instead of occurring princi- y in the middle class.

he score method of class allocation was again applied and in this gave a higher level of agreement with interviewer judgements than the discriminants, suggesting that the theoretical model was not ropriate. A comparison of the results of the three classifications is n in Table X.

TABLE X

Levels of agreement between different classificatory methods

% Agreement between:	Males	Females
Scores and interviewers	72	77
Discriminants and interviewers ..	72	69
Discriminants and scores	83	82

ne results obtained from the random and quota samples are very ar. Neither the different methods of sampling nor the use of viewers from several organisations appear to have had any effect he overall level of agreement between discriminants and inter- ers.

constructing discriminant functions based on attributes other than employed by us in this study, a closer approximation to the inter- ers' classification might be obtained. It seems unlikely, however,

that any very high level of agreement could be realised, for th appear to be no other completely objective attributes more clo related to social class than those used; and if subjective factors such accent and appearance were taken into account, variability betw bases of evaluation of different interviewers would certainly incre

The results of this pilot study do not necessarily prove the int viewers to have been inconsistent in their assessments of social cl but they certainly help to substantiate the widespread fear that su inconsistency must exist. The discriminant classification can, howev only be considered better than that of the interviewers, on which it w based, in that it is completely objective.

In order to obtain more satisfactory discriminant functions, vario modifications of our method are possible. In the first place, the samp of informants could be classified in some way other than by an ordina interviewer. Classification could be carried out by a panel of suitab qualified persons, or by some strictly objective method based on t answers of these informants to a detailed questionnaire relating to the spending and leisure habits.

If a suitably classified sample had been obtained, a more refin method could be used in the scoring of attributes. A 'normal score' the type used by Dennis Chapman in his recent study of social stat would ensure that the assumption of normality in the distributions attributes did correspond to the figures. If discriminants based on su data were then found to classify with reasonable accuracy, it seems t they would provide a suitable objective measure of socio-econor class which would be of real value in market research. With this mind, a project is now under consideration by the Research Techniq Unit and may soon be undertaken.

REFERENCES

[1] Barnard, M. M. (1935). 'The secular variations in skull characteristics in f series of Egyptian skulls', *Annals of Eugenics*, **6**, 352.

[2] Wallace, N. and Travers, R. M. W. (1938). 'A psychometric sociological st of a group of speciality salesmen', *Annals of Eugenics*, **8**, 266.

[3] Mannheim, H. and Wilkins, L. T. (1955). *Prediction Methods in Relation to Bo Training*. HMSO, London.

[4] Moser, C. A. and Stuart, A. (1953). 'An experimental study of quota samplin *J. R. Statist. Soc.*, A, **116**, 23.

[5] Dwyer, P. S. (1951). *Linear Computations*. Wiley, New York.

[6] Rao, C. R. (1952). *Advanced Statistical Methods in Biometric Research*. Wiley, N York.

[7] Cauter, T. and Downham, J. S. (1954). *The Communication of Ideas*. Chatto a Windus, London.

[8] Chapman, D. (1955). *The Home and Social Status*. Routledge and Kegan Pa London.

THE STUDY OF VARIABILITY IN ENGINEERING DESIGN

S. JAMES MORRISON

The British Thomson-Houston Co. Ltd., Rugby

Mr Morrison shows how a fairly straightforward application of statistical theory can help the designer. A rough prediction of the variability of a product can be made, and the elements in the design that are likely to give the most trouble can often be identified.

duction

ractical applications of statistical theory are frequently met in e-scale manufacture (e.g. statistical quality control), but examples ne earlier stages of engineering design and development are less mon. This is unfortunate, since manufacturing problems can often nticipated and allowances made for known variables if the design rmulated on a statistical basis. Such a design is more likely to eed in production, with less risk of time being wasted on investi- ng snags and money being wasted on design changes after produc- has commenced.

Iost engineering design is based to a large extent on relations be- en mean values or target values of the design parameters. For a stical formulation each equation of mean values should be supple- ted by an equation in terms of the variance of the design para- ers. This will require a knowledge of the component variances, ch should be based on actual measurements rather than on guess- k. The equation of variance can sometimes be derived directly a the equation of means. The value of the variance equation to the gner is twofold: (i) it will enable the designer to base an estimate e overall variance of a new design on his experience of the variance ne component parameters in earlier designs that have reached the luction stage; (ii) it may guide the designer towards better propor- s or a better arrangement in a design, so as to reduce the overall ance.

he variance equation is also useful in studying manufacturing blems, since it will show which factors in a complex situation are ing large contributions to the overall variance.

he examples given in this article are taken from the specialised of glass engineering design, and are concerned with the manufac- of glass beads and rings used as components in vacuum envelopes glass-metal seals for electronic valves and semiconductor devices.

There is no reason why similar methods should not be used m widely in engineering design. The method employs a stand statistical technique, but so far as is known it has not received m attention in this field.

Theory

The following symbols will be used for the dimensions of beads a rings.

D = outside diameter

B = bore (B_1, B_2, B_3, etc., in multi-bore beads)

n = number of bores

w = wall thickness (of rings)

L = length

W = weight

G_1 = volume determined by weighing

G_2 = volume estimated from dimensions

ρ = density

Mean values will be represented by $\bar{D}$, $\bar{B}$, etc., and variances will represented by $V(D)$, $V(B)$, etc. It will be assumed that the dimensio are independent variables.

It is shown in several text-books[1, 2] that if X is a general function independent variates,

$$X = f(x_1, x_2, x_3, \ldots, x_n)$$

then the variance of X is given by

$$V(X) \simeq \left(\frac{\partial X}{\partial x_1}\right)^2 . V(x_1) + \left(\frac{\partial X}{\partial x_2}\right)^2 . V(x_2) + \left(\frac{\partial X}{\partial x_3}\right)^2 . V(x_3) + \ldots + \left(\frac{\partial X}{\partial x_n}\right)^2 . V(x_n) \quad \ldots . ($$

The formula is exact for linear functions but should only be appli to non-linear functions if the standard deviation of each variable x less than 20 per cent of the mean.

In all cases the volume determined by weighing is given by

$$G_1 = \frac{W}{\rho} \quad \ldots . (2$$

The theoretical treatment of different types of bead will vary.

(*a*) *Single-hole beads.* The estimated volume is given by

$$G_2 = \frac{\pi}{4}(D^2 - B^2)L \quad \ldots . (3$$

m (1) and (3) the variance equation is

$$V(G_2) = \left[\frac{\pi DL}{2}\right]^2 . V(D) + \left[\frac{\pi BL}{2}\right]^2 . V(B) + \left[\frac{\pi}{4}(D^2 - B^2)\right]^2 . V(L) \quad \ldots . (4)$$

Rings. The estimated volume is given by

$$G_2 = \pi wL(D - w) \quad \ldots . (5)$$

m (1) and (5) the variance equation is

$$V(G_2) = [\pi wL]^2 . V(D) + [\pi w(D - w)]^2 . V(L) + [\pi L(D - 2w)]^2 . V(w) \quad \ldots . (6)$$

Multiple-bore Beads; Unequal Bores. The estimated volume is given

$$G_2 = \frac{\pi}{4}[D^2 - (B_1^2 + B_2^2 + B_3^2 + \ldots + B_n^2)]L \quad \ldots . (7)$$

m (1) and (7) the variance equation is

$$V(G_2) = \left\{\frac{\pi DL}{2}\right\}^2 . V(D) + \left\{\frac{\pi}{4}[D^2 - (B_1^2 + B_2^2 + B_3^2 + \ldots + B_n^2)]\right\}^2 . V(L) + \left\{\frac{\pi B_1 L}{2}\right\}^2 . V(B_1) + \left\{\frac{\pi B_2 L}{2}\right\}^2 . V(B_2) + \left\{\frac{\pi B_3 L}{2}\right\}^2 . V(B_3) + \ldots + \left\{\frac{\pi B_n L}{2}\right\}^2 . V(B_n) \quad \ldots . (8)$$

Multiple-bore Beads; Bores Nominally Equal. The estimated volume n individual bead is given by equation (7) above. The variance ation can be derived from equation (8) by putting $B_1 = B_2 = B_3 \ldots = B_n = B$, thus:

$$V(G_2) = \left\{\frac{\pi DL}{2}\right\}^2 . V(D) + \left\{\frac{\pi}{4}[D^2 - nB^2]\right\}^2 . V(L) + n\left\{\frac{\pi BL}{2}\right\}^2 . V(B) \quad \ldots . (9)$$

Multiple-bore Beads; Bores Exactly Equal. In this case equation (7) omes

$$G_2 = \frac{\pi}{4}[D^2 - nB^2]L \quad \ldots . (10)$$

From (1) and (10) the variance equation is

$$V(G_2) = \left\{\frac{\pi DL}{2}\right\}^2 . \mathrm{V}(D) + \left\{\frac{\pi}{4}[D^2 - nB^2]\right\}^2 . \mathrm{V}(L) + \left\{\frac{n\pi BL}{2}\right\}^2 . \mathrm{V}(I$$

....(11)

This is similar to equation (9), but the component of variance due t bore is n times greater.

In the present case the method was tested by calculating the mea and variance of the volume G_2 from the appropriate equations b substituting the measured values of the mean and variance of the dimensions. The use of the *arithmetic mean* of dimensions is not strictl correct, but when the standard deviation of a dimension is less tha 20 per cent of its mean the error so introduced will be small.

Measurements

The theoretical results given above have been tested by measurements on groups of glass beads and rings of four different types. Diameter, bore (or wall thickness), and length were measured with a measuring microscope, and weight was measured on a torsion balance. The volume was calculated from the weight and the known value o the density. The mean and the variance of dimensions were calculate for each group of beads.

For all but one of the groups the mean volume $\bar{G}_1$ determined b weighing was found to be in close agreement with the mean volume $\bar{G}$ calculated from mean dimensions. The exception was a group of thirt triple-bore beads, for which a difference significant at the 10 per cen level was found. The beads were re-examined and it was found tha outside diameter and bore were non-circular to a marked extent owing to a defect in manufacture.

The volume variance $\mathrm{V}(G_1)$ determined by weighing was compare with the estimated variance $\mathrm{V}(G_2)$ calculated by the appropriat formula, (4), (6), (8), (9), or (11). The results are summarised i Table I. These results show that the variance of the volume of glas beads and rings can be calculated with a fair degree of accuracy fro the dimensional variances, and that any errors are likely to be on the safe side (i.e. the actual variance will usually be less than the estimated variance). If there is any significance at all in the discrepancies it is probably due to a slight degree of correlation between bore and outside diameter in the tube-drawing process and to the errors of measurement. All the beads and rings were made by cutting from hand-drawn tubing, which is characterised by considerable dimensional variation. The measurements will therefore be subject to errors associated with ovality and with variation of wall thickness. There was also some degree of chipping, which would affect the cut length. These errors would swell the $\mathrm{V}(G_2)$ values more than the $\mathrm{V}(G_1)$ values, which depend on only one type of measurement. The variability of density, which is small relative to the dimensional variation, has been ignored in calculating $\mathrm{V}(G_1)$.

TABLE I

Dimensional variance of glass beads (all dimensions in mm.)

	Standard-wall Bead	Thick-wall Bead	Thin-wall Ring	Triple-bore Bead	
mber of beads ..	30	30	36	30	
tside meter $\bar{D}$ ..	1·69	5·12	7·09	2·22	
$V(D)$ ..	0·001 25	0·010 0	0·061 9	0·022 0	
re $\bar{B}$ ..	0·625	2·29	—	0·559	
$V(B)$ ..	0·002 54	0·004 32	—	0·006 59	
all ckness $\bar{w}$ ..	—	—	1·02	—	
$V(w)$ ..	—	—	0·016 2	—	
ngth $\bar{L}$ ..	1·92	2·33	12·51	2·22	
$V(L)$ ..	0·005 36	0·027 9	0·102 0	0·024 2	
lume ermined weighing $\bar{G}_1$ ..	3·70	38·2	242·4	6·64	
$V(G_1)$ ..	0·038 8	7·378	873	0·936	
imated ume ote 1) $\bar{G}_2$ ..	3·72	38·4	243·3	6·96	
$V(G_2)$ ..	0·061 7	11·39	775	1·628	1·778
riance equation ..	(4)	(4)	(6)	(9)	(11)
riance ratio F .. .. otes 2 *and* 3)	1·59	1·54	1·13	1·74	1·90
obability of exceeding ratio	$P \simeq 20\%$	$P \simeq 20\%$	$P \geqslant 50\%$	$20\% > P > 10\%$	$10\% > P > 5\%$
vel of significance ..	Low	Low	Low	Low	Borderline

e 1. The degrees of freedom for the estimated volume have been taken to be the same the volume determined by weighing. According to Note 2 the degrees of freedom be greater, but the exact values are unimportant, since F is not very sensitive to small es in degrees of freedom.

e 2. Strictly speaking, the F-test is not valid in this situation, since the estimated ce given by equation (1) is not distributed as χ^2, and so the ratio of the estimated ce to the measured variance is not distributed as F. An approximate method for g with this situation, and for determining the degrees of freedom of the composite mean , has been described by Le Roy P. V. Johnson and Ernest S. Keeping.[3] This method give levels of significance slightly different from those in the Table, but the main sions would not be affected.

3. Since differences in variances are being compared in both directions, a two-tail is appropriate.

e individual terms in the variance equations show how much of otal variance of the volume can be ascribed to each dimensional nce. This analysis, given in Table II, shows that in each type of there is a predominant factor which makes a large contribution e final volume variance, and which varies between one type of and another. In situations in which bead volume is a critical (e.g. in the manufacture of glass-metal seals) the designer can is analysis as a guide in determining the most suitable proportions

TABLE II

Estimated variance of the volume of glass beads apportioned among different sources

	Standard-wall Bead	Thick-wall Bead	Thin-wall Ring	Triple-bore Bead	
Equation	(4)	(4)	(6)	(9)	(11)
Source:					
Diameter	0·032 6	3·52	99	1·315	1·315
Bore	0·009 0	0·30	—	0·075	0·225
Wall	—	—	637	—	—
Length	0·020 1	7·57	39	0·238	0·238
Total variance	0·061 7	11·39	775	1·628	1·778

of bead for a given application, and can compare the effect of differe processes of bead manufacture which have different degrees of dime sional accuracy. The analysis also gives the manufacturer a valua guide, pointing to the most likely source of trouble if the overall proc variation is too large.

It is reasonable to suppose that a similar statistical analysis mig be applied in other fields of engineering design and manufacture a might lead to equally useful results.

Acknowledgements

I wish to acknowledge the assistance of colleagues in the BTH Resea Laboratory, in particular that of Mr G. R. Lines, for his painstak measurements on glass beads. I am also grateful to Mr G. B. Redmo of the Dunlop Research Centre, and to Mr E. D. van Rest, for s gestions and comments.

REFERENCES

[1] Davies, O. L., Editor (1947). *Statistical Methods in Research and Production*, p. Oliver and Boyd, Edinburgh and London.

[2] Tippett, L. H. C. (1952). *The Methods of Statistics*, p. 124. Williams and Norg London; Wiley, New York.

[3] Johnson, L. P. V. and Keeping, E. S. (1952). 'Composite mean squares and t degrees of freedom', *Applied Statistics*, **1**, 202.

A SAMPLE SURVEY WITH BOTH A POSTAL AND AN INTERVIEW STAGE

PERCY G. GRAY
Government Social Survey

In this article Mr Gray shows how a sample survey dealing with the general public was carried out to a large extent by means of a postal inquiry, with satisfactory results. He suggests that the postal approach is well worth consideration for certain types of survey problem.

main problem in a postal inquiry is to get a good response, and neral the Social Survey has carried out relatively few inquiries by Where such a postal approach has been used in the past, the ling population has always been a specialised one: midwives, uates, doctors, and so on, who might be expected to be particularly ested in the subject of the inquiry and who might, moreover, be osed to be rather more capable of dealing with forms. With such alised populations it seems possible to get a response rate solely by of over 80% after two reminders and using a fairly lengthy ques- aire. Until the inquiry now to be described our experience with eneral population was limited to people who had previously been viewed and who were then sent a simple questionnaire by post. a questionnaire dealing with medical consultations we obtained onse of 80% without any reminder letter. With one dealing with acle-wearing we obtained a response of 95% after using a reminder . At the time, we attributed these good results to the interest and will aroused by the previous interview rather than to the simplicity e questionnaires. The results now to be presented suggest that the response was due to the simplicity of the questionnaires, since in case there had been no previous interview. While there are ably very few surveys as simple as this one in which the full ionnaire can be administered by post, the results are of interest they suggest a means of securing samples of some of those thinly red groups of people that trouble the sampler.

e inquiry[1] to be described was made in November and December for the Ministry of Agriculture and Fisheries. Its object was to ate the numbers of poultry and pigs in England and Wales, ding two groups. The first group to be excluded consisted of kept on agricultural holdings of more than one acre in extent and covered by the quarterly Agricultural Censuses. The second to be excluded consisted of those kept on under-one-acre holdings

where the keeper obtained the feeding-stuffs rations for his livestock June 1953 direct from the County Agricultural Executive Committ such holdings appeared on a list held by the Ministry and were be approached separately by the Ministry. It was known that an app ciable number of pigs and poultry were kept in ordinary gardens and other holdings not covered by these two groups and, particularly af the end of feeding-stuffs rationing in August 1953, very little was kno about their numbers. This was the group for which we were asked provide estimates. Since no records of these poultry- and pig-keepe were available, we had to design a sample covering all household Our main objective was to determine the number of fowls, and it w with this statistic in mind that we designed the inquiry.

Examination of some data obtained from another inquiry enabled to see how the proportion of households keeping poultry varied wi the type of area. They showed that whereas only about 10% of households kept non-farm poultry, about 30% of those living in rur districts did so. Furthermore, in urban areas the proportion of hous holds keeping poultry decreased markedly with increasing populati density. The data suggested that only about 5% of the poultry wou be found in urban wards with a population density of ten or mo persons per acre. We decided therefore to divide the country into th strata:

(1) Rural Districts

(2) Urban wards with a density of less than ten persons per acr

(3) Other urban wards

and to use a variable sampling fraction. Examination of the roo estimates of the variances which were available showed, however, th even with the most efficient distribution of the sample a large num of households would need to be approached to give the accuracy ask for.

Faced with the need for such a large sample, and bearing in m the simple factual nature of the information required from each hou hold, we decided to investigate the possibility of doing the inquiry post. As a first step we sent out a pilot questionnaire to a sample 100 rural addresses. After about a week we sent a reminder letter those who had not replied and by the end of two weeks had obtain a total of 79 questionnaires. Encouraged by this, we repeated th experiment for the low-density urban stratum and obtained rath similar results. These results suggested that the foregoing two stra might be dealt with by sending out postal questionnaires and th interviewing a sub-sample of the non-respondents. Cost considerati suggested that only half of the non-respondents to the postal questio naires needed to be interviewed. We were hesitant, however, to addr a postal questionnaire on poultry- and pig-keeping to household living in the more densely populated urban areas. Fortunately

able to obtain some data for this stratum by adding some of the ions to another survey covering a national sample of households. e considerations determined the plan of the inquiry, which is set n Tables I and II.

TABLE I

The sample design

	(1) Rural Stratum	(2) Low-density Urban Stratum	(3) High-density Urban Stratum
-stage units	50 *Administrative Rural Districts* selected with probability proportional to population	40 *urban wards with density of less than* 10 *persons per acre* selected with probability proportional to number of structurally separate dwellings occupied	Estimates for this stratum were based on information obtained by putting extra questions to the relevant parts of a national sample of households interviewed for another inquiry carried out at about the same time. The relevant part of the design consisted of the households living at 20 addresses selected from the Electoral Register in each of 21 *urban wards with densities of* 10 *or more persons per acre*, a total of 420 addresses.
nd-stage units	Separately rated items of dwelling property selected at a constant interval from the rate-books of the Local Authority	Addresses selected at a constant interval from the Electoral Register	
ıber of second- units per first- unit	120	100	
l second-stage	6 000	4 000	

will be seen that in both the rural and the low-density urban a the sample design was of the two-stage type. This is the part of ample with which this article is concerned. In the selected rural cts the second stage of sampling was done from the rate-books, eas in the selected low-density urban wards addresses were selected the current Electoral Register. We used the rate-books for the districts because they provide a rather better sampling frame for ing rural addresses when the approach is to be made by post. nally in an interview survey we can rely on the interviewer to ify the selected unit on the spot. A number of operations had to rried out on the resulting sample-lists before the questionnaires be sent out.

vestigators selecting the samples from rate-books in rural districts instructed to record for each selected item of property the brief iption of the property given in the rate-book, the name of the oier (where this was given), the full address of the property, its ble value, and its assessment number. When each list was received adquarters, items which clearly contained no living accommoda- (lock-up shops, etc.) were excluded, and any surplus of the other

items above the 120 required was rejected at a constant interval. T
remainder were serial-numbered 1 to 120 throughout the list. T
addresses of the serial-numbered items were then further scrutinis
any which had no name of occupier and an inadequate postal addr
were specially marked for exclusion from the postal stage of the inqui

The samples of addresses drawn from the Electoral Register in t
low-density urban wards (which included, by definition, hardly a
addresses without residential accommodation) were reduced to t
required size of 100 in a similar way and were each serial-numbe
1 to 100.

The addresses in the sample which we could at this stage identi
with certainty as farmhouses were then withdrawn, since by definiti
they could contain no poultry or pigs in our categories.

A postal questionnaire was sent out on 17th November to th
remainder of the sample, that is to all except those with inadequa
postal addresses or where the address was known to be a farm. Th
questionnaire* consisted of a single printed sheet $6\frac{1}{2} \times 8$ in. bearin
nine questions, with provision for writing an overriding 'NONE' if n
poultry or pigs had been kept at the address. The covering lette
consisted of a single-sided sheet of the same size addressed to 'De
Householder', and explained the purpose of the inquiry. All instru
tions were on the questionnaire, but the letter did stress that the for
should be returned even if people 'have not kept any hens or pigs a
indeed have no opportunity for doing so'. A reply-paid envelope w
enclosed. The informants were not asked to give their name a
address, though many did so. Identification was by a serial numb
added to the questionnaires. In a future inquiry it might be wo
adding the serial number to the envelope, since a small number
people returned the covering letter with the word 'none' written on
but without the questionnaire. Fortunately in most cases they add
their address.

In those cases where a reply had not been received by the afterno
post of 25th November a reminder letter was despatched on that da
A questionnaire returned by the Post Office saying definitely that t
premises were unoccupied was treated as a return; but where t
reason for non-delivery was given as 'address untraceable', 'unknow
etc., the address was retained for further follow-up at the interviewin
stage. The reminder letter used was very short, consisting only of tw
paragraphs: we did not enclose a further questionnaire.

After 6th December the odd and even serial-numbered address
were treated in different ways as will be seen from Table II, whic
summarises the sequence of operations.

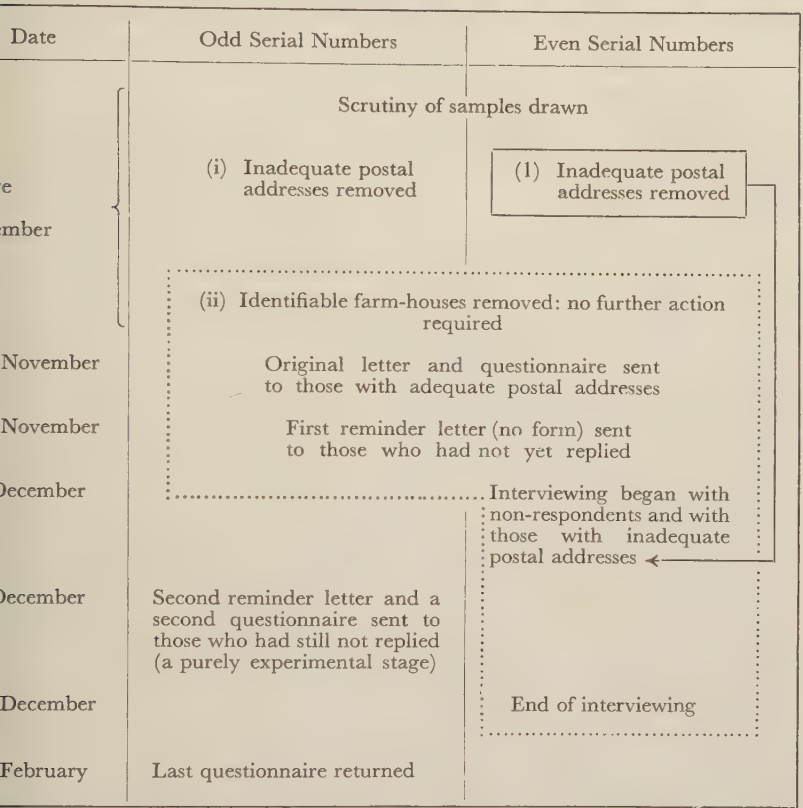
It will be noted that the sub-sampling of the non-respondents fo
follow-up by interviewers was, for convenience, done on the basis
the serial numbers originally allocated to the selected addresses: on

* Since the layout and wording in postal inquiries may well be critical, the postal questio
naire and letters are reproduced in the Appendix.

with an even serial number were followed up by interviews. a sample of approximately, though not exactly, one half of those had not replied by ten days after the first reminder letter had been out were followed up in this way. In addition, members of the

TABLE II

ce of operations on the sample of 10 000 *drawn in the rural and low-density urban strata*


Date
Odd Serial Numbers
Even Serial Numbers
Scrutiny of samples drawn
(i) Inadequate postal addresses removed
(1) Inadequate postal addresses removed
(ii) Identifiable farm-houses removed: no further action required
November
Original letter and questionnaire sent to those with adequate postal addresses
November
First reminder letter (no form) sent to those who had not yet replied
December
Interviewing began with non-respondents and with those with inadequate postal addresses
December
Second reminder letter and a second questionnaire sent to those who had still not replied (a purely experimental stage)
December
End of interviewing
February
Last questionnaire returned

E. Estimates made in the report were based only on the operations enclosed within ted line.

le which had been excluded from the postal stage because of an quate postal address were similarly sub-sampled, and those with serial numbers were issued to interviewers along with the non-ndents. While the postal stage was going on, opportunity had taken to seek further information about this sub-sample of in-uate addresses from the rating officers concerned, and we were able to give the interviewers extra information about the majority m.

e lists of addresses were sent out to the interviewers on Saturday, ecember, but any questionnaires received by us on Monday, 6th nber, were treated as postal returns, since they could not have affected by the interviewing. From then onwards any even-serial-

numbered questionnaires received by post were treated as interv
returns. Any odd-numbered questionnaires received after this c

TABLE III

The response

Postal Results		Rural Stratum				Low-density Urban Stratum			
		Odd Serial Numbers		Even Serial Numbers		Odd Serial Numbers		Even Serial Numbers	
Date Received		No.	Cumulative %	No.	Cumulative %	No.	Cumulative %	No.	Cumu tiv
Friday	Nov. 19	257	9·2	220	7·8	289	14·5	285	14
Saturday	Nov. 20	405	23·6	430	23·0	331	31·1	314	30
Monday	Nov. 22	220	31·5	258	32·1	141	38·2	154	37
Tuesday	Nov. 23	174	37·7	171	38·1	115	44·0	109	43
Wednesday	Nov. 24	102	41·3	119	42·3	61	47·0	83	47
Thursday	Nov. 25	96	44·8	71	44·9	64	50·3	58	50
Friday	Nov. 26	60	46·9	68	47·3	47	52·6	35	52
Saturday	Nov. 27*	229	55·1	209	54·6	157	60·5	143	59
Monday	Nov. 29	281	65·1	261	63·9	214	71·2	198	69
Tuesday	Nov. 30	176	71·4	187	70·5	104	76·5	107	74
Wednesday	Dec. 1	108	75·2	111	74·4	56	79·3	57	77
Thursday	Dec. 2	54	77·2	58	76·5	21	80·3	42	79
Friday	Dec. 3	28	78·2	33	77·6	25	81·6	20	80
Saturday	Dec. 4	23	79·0	25	78·5	17	82·4	14	81
Monday	Dec. 6	18	79·6	17	79·1	9	82·9	12	81
Tuesday	Dec. 7	10	80·0			5	83·1		
Wednesday	Dec. 8	7	80·2	Any postal		7	83·5	Any postal	
Thursday	Dec. 9	4	80·4	returns after		2	83·6	returns after	
Friday	Dec. 10†	69	82·8	6th December		56	86·4	6th Decemb	
Saturday	Dec. 11	83	85·8	have been		71	90·0	have been	
Monday	Dec. 13	46	87·4	credited to the		19	90·9	credited to t	
Tuesday	Dec. 14	44	89·0	interview stage.		26	92·2	interview sta	
Wednesday	Dec. 15	28	90·0	The second		15	93·0	The second	
Thursday	Dec. 16	4	90·1	reminder was		7	93·3	reminder wa	
Friday	Dec. 17	8	90·4	not sent to		3	93·5	not sent to	
Saturday	Dec. 18	1	90·5	this half of		1	93·5	this half of	
After 18th Dec.		36	91·8	the sample.		14	94·2	the sample.	
Approached but not returned	(*a*)	231		591		115		362	
Total approached by post	(*b*)	2 802	100·0	2 829	100·0	1 992	100·0	1 993	10
Inadequate postal address ..	(*c*)	132		105		2		2	
Unnecessary to approach ..	(*d*)	66		66		6		5	
		3 000		3 000		2 000		2 000	
Interview Stage									
Accounted for (*e*)				682				360	
Not accounted for (*f*):									
Refusal				4				–	
Address not visited				4				3	
Omitted in error				6				1	
Eligible for this stage (*g*)				696				364	
Overall Result			%		%		%		%
Total accounted for (*h*)		2 637	87·9	2 986	99·5	1 883	94·1	1 996	99
Total unaccounted for (*j*)		363	12·1	14	0·5	117	5·9	4	0
Total Sample		3 000	100·0	3 000	100·0	2 000	100·0	2 000	100

Notes. (*g*) = (*a*) + (*c*); (*h*) includes (*d*).

* Effects of first reminder letter. † Effects of second reminder letter.

kept separate and did not enter into the estimates made. Thus estimates were based on both odd and even questionnaires up to December but only on even-numbered ones after that date. The had, of course, to be re-weighted.

owever, we continued to receive odd-numbered schedules and on December sent out a second reminder letter to those addresses with serial numbers from which no reply had been received. This was xperimental stage. We included with this letter a further copy of questionnaire which could readily be distinguished from that nally sent. Interviewing at the even-numbered addresses ceased 8th December, but postal questionnaires continued to come in odd-numbered addresses for some time, one being received on February.

etails of the response obtained both by post and interview are given able III. The top section of the table gives the day-by-day response e postal questionnaire by the two rural samples and by the two density urban samples. The cumulative percentages given are l on the number of addresses approached by post, not on the full le. For example, in the sample formed by the odd-serial-numbered addresses there were 66 addresses which were known to be farms which therefore did not need to be approached either by post or view. There were a further 132 addresses which were not written we had some doubt as to their adequacy for postal purposes. This ber (4% of the rural addresses) may seem unduly high, but it ded all cases where we had any doubts about the questionnaire ing the correct address. Difficulty mainly arises with farms and cottages.

e first day's post brought us 9·2% and 7·8% of the two rural les and 14·5% and 14·3% of the low-density urban samples. A later, on the day before the first reminder took effect, these per-ges had risen to 46·9% and 47·3% for the rural samples and % and 52·1% for the others. The first reminder letter brought er flood of replies, and by 6th December, when the interviewing nenced with the even serial numbers, the proportions had reached % and 79·1% for the rural samples and 82·9% and 81·8% for the n samples.

th the odd-serial-numbered samples the second reminder letter copy of the questionnaire brought yet another burst of answers, he proportions finally reached 91·8% for the rural sample and % for the urban sample.

e results for the two entirely postal samples are presented in a y different form in Fig. 1. The heavy line on the graphs gives roportion of those approached by post who replied on successive up to the day before the first reminder took effect. The broken ives the proportion replying on successive days, starting with the hen the first reminder took effect and continuing up to the day the second reminder took effect. All the percentages represented

by this line are based on the number of people who had not repl prior to the first reminder taking effect. Similarly the percenta represented by the dotted line are based on the number of people v had not replied prior to the second reminder taking effect. This man

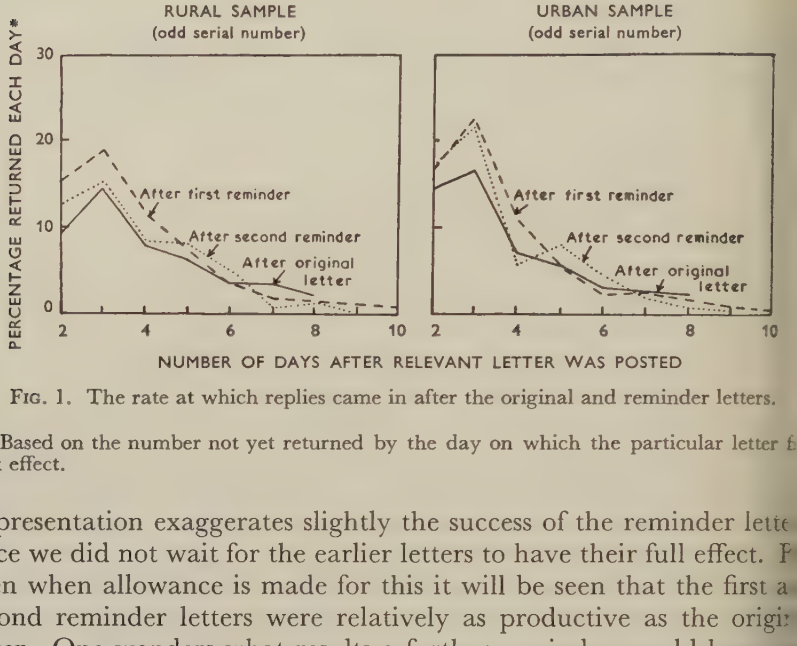


FIG. 1. The rate at which replies came in after the original and reminder letters.

* Based on the number not yet returned by the day on which the particular letter t took effect.

of presentation exaggerates slightly the success of the reminder lett since we did not wait for the earlier letters to have their full effect. F even when allowance is made for this it will be seen that the first a second reminder letters were relatively as productive as the origi letter. One wonders what results a further reminder would have p duced. The rate at which the returns came in suggests that in such inquiry the first reminder should be sent seven days after the origi letter and the second reminder seven days after that.

In addition to giving the results for the postal stage Table III gi the results for the interview stage and the overall results. In the c of the even-serial-number rural sample the interviewers were requi to visit the 591 addresses from which a reply had not been received a also the 105 addresses which we had originally considered inadequ for postal purposes (we had by this time secured fuller addresses fr the rating authorities and the Electoral Roll). At 372 (63%) of the 5 addresses to which we had sent postal questionnaires our informa generally the housewife, remembered the arrival of the form and in (16%) cases was able to produce it. Thus in about 4 out of 5 cases t questionnaire was by that time probably lost or destroyed. This su gests that we were fully justified in sending a further copy of t questionnaire with the second reminder letter to the completely post samples. It may be noted that after 6th December we received 3 replies from the completely postal rural sample, 55 on the origin questionnaire and 285 on the copies.

Of the 696 addresses eligible for interview in the rural even-seri numbered sample we accounted for 682. There were 4 refusals, 4 cas

ere the interviewers failed to visit an address, and 6 cases where we ed to issue the addresses to the interviewers. With the urban sample accounted for 360 out of the 364 addresses: there were no refusals. The overall results are given at the bottom of Table III. For the npletely postal rural sample 87·9% of the addresses were accounted compared with 99·5% for the rural sample with interviewing. The responding proportions for the urban samples were 94·1% and 3%. This comparison is perhaps rather unfair to the completely tal rural sample, since no attempt was made to deal with the 132 dequate postal addresses originally found in the sample. If these had n treated in the same way as those issued to the interviewers and itional information obtained, it would certainly have been possible vrite to most of them and the response would have approached 92%. nerally it will be seen that the proportion of addresses accounted was gratifyingly high, in the case of the postal samples surprisingly

TABLE IV

The variation in the proportion of respondents answering 'None' (rural stratum only)

	Odd Serial Numbers		Even Serial Numbers	
	%		%	
First day's post	86	(257)*	82	(220)*
Second day's post	77	(405)	73	(430)
Third day's post	70	(220)	72	(258)
Remaining post until reminder took effect ..	58	(432)	62	(429)
fter first reminder:				
First day's post	61	(229)	65	(209)
Second day's post	59	(281)	60	(261)
Third day's post	57	(176)	57	(187)
Remaining post before next stage	57	(252)	58	(244)
fter second reminder:				
First day's post	67	(69)	—	
Second day's post	60	(83)	—	
Third day's post	54	(46)	—	
Remaining post	57	(121)	—	
p to first reminder	71	(1 314)	71	(1 337)
p to second reminder ..	66	(2 252)	—	
l postal returns	65	(2 571)	—	
terviewees who had been approached by post ..	—		61	(577)
l possible postal returns† ..	—		65	(2 815)

Numbers in brackets are the bases of the percentages.
Except for the 14 cases where we failed to get any information (see Table III).

Ve have now to consider the bias that would have resulted if we had n the first day's post only; if we had sent no reminder; and so on. llustrate these points the results for the rural stratum alone will be . They are best demonstrated by considering the proportion of

respondents who could simply answer 'none': that is, they had had pigs or poultry in the period with which we were concerned, nor w they farmers. How this proportion varies is shown in Table IV.

As a yardstick for measuring the bias we have the result obtained the even-serial-numbered sample, where after the interview stage had accounted for all but 14 of those approached by post. The pr portion answering 'none' based on this sample of 2 815 is 65%.

Clearly an estimate based on the first day's post would be bad biased. Ease of filling in the form influences both the speed and chan of the form being returned. Successive posts, up to the first remind taking effect, produced a higher proportion of livestock keepers, i. those who would have to expend more time and effort in completin the form. The first post after the first reminder letter took effect pr duced a temporary increase in the proportion answering 'none', bu this fell away with succeeding posts. The second reminder letter (wit another form) produced similar results.

We see that if we had sent no reminder letter we should have pr duced an over-estimate (71%). If we had stopped after the fir reminder we should have produced quite a good estimate (66%. However, if we had decided to use the three postal stages we shou have obtained a figure of 65%, which agrees exactly with that obtain after the almost complete follow-up of all those approached by po (Actually the percentage is 65·4% in both cases.)

It will be remembered that a group of addresses were not approach by post because the addresses were originally considered inadequa However, by the time we issued them to the interviewers we h obtained further information, and it now seems certain that most these cases could then have been successfully approached by po Even if this group is not allowed for, we have found no measurable bi in an estimate based solely on the odd-serial-numbered postal retur for the most important item in the survey. This item is the avera number of domestic fowls per address at the time of the survey. (It d not include fowls kept on farms or those kept by certain keepers on t Ministry lists.) Based on the 2 637 odd-serial-numbered addres accounted for by post alone, this average is 2·72; based on the 2 9 even-serial-numbered addresses accounted for by post and intervie this average is 2·77. The difference between these two averages is n statistically significant, so we may conclude that there would be n measurable bias in an estimate based solely on the postal returns.

It then becomes clear that this would have been the most efficie way to carry out the inquiry, that is, entirely by post. Because we ha planned for an interview stage our sample design was of two stages, th first stage consisting of 50 rural districts (out of 477). With this tw stage design our sampling standard error for the average given abo was 0·23 as compared with the value of 0·14 which we could expe from a sample of the same size covering all rural districts. If we had n planned an interview stage we should not have needed to concentra

sample in 50 first-stage units, but could have spread it over a er number of rural districts. Thus had we relied solely on the we should have produced no measurable bias, and by diverting resources from interviewing to increasing the number of rural ricts covered we should have considerably reduced the sampling r.

he fact that a completely postal approach produced no measurable in this case is of course no guarantee that a bias would not arise in er cases. Some follow-up of those who do not reply to the postal stionnaire will probably be necessary. This immediately raises the stion of what is the optimum fraction for sub-sampling the non-ondents. Cochran[2] has shown that this is given by

$$K_{\text{opt}} = \sqrt{\left[\frac{C_2}{C_0}\left\{W_1 + \left(\frac{S^2}{S_2^2} - 1\right)\right\} \Big/ \left\{1 + \frac{C_1}{C_0}W_1\right\}\right]}$$

re W_1 = the proportion responding by post

S^2 = the variance for the whole population, for the most important item in the inquiry

S_2^2 = the variance for those not responding by post

C_0 = the cost per unit of preparing the sample and the postal stage

C_1 = the processing cost per unit

C_2 = the cost for interviewing and processing one unit

will be remembered that in planning the inquiry K was taken to 2, that is, half the non-respondents to the original letter and the reminder were approached by the interviewers. After the event it ossible to calculate K_{opt} more accurately. For the rural stratum we that $S^2/S_2^2 \simeq 1$, $C_2/C_0 \simeq 9{\cdot}5$, $C_1/C_0 \simeq 0{\cdot}15$, and $W_1 \simeq 0{\cdot}8$, giving = 2·6. It seems therefore that we might perhaps have taken = 3 instead of $K = 2$. Examination of the expression for K_{opt} given ve suggests that with values of C_2/C_0 likely to be found in surveys ing with individuals or households K_{opt} is unlikely to exceed 3 ss S^2/S_2^2 exceeds unity by an appreciable amount. It is, however, difficult to measure this in advance from the results of a small t, and the assumption has usually to be made that $S^2/S_2^2 = 1$. Thus sampling fractions of $\frac{1}{2}$ or $\frac{1}{3}$ are likely to cover most cases.

erhaps I should end this article by stressing the simple nature of the stionnaire used in the inquiry I have described, for I suspect that presence of a single question that is difficult to answer might very h alter the response. We would not suggest for one moment that y surveys could be done entirely by post. However, these results e made us reconsider the possibility of using the postal approach as ethod of locating relatively scattered sampling populations for ch lists do not exist. A carefully framed postal questionnaire might

serve to locate such a population.* Interviews would then be ma with the respondents having the required characteristics together w a sub-sample of the non-respondents. If a response rate similar to th on the inquiry described can be achieved such an approach becom well worth while.

Finally I should like to thank my former colleague, Tom Corle with whom I had the pleasure of working on this inquiry, and t Ministry of Agriculture, Fisheries and Food for permission to publis this article.

REFERENCES

[1] Gray, P. G. and Corlett, T. (1955). Social Survey Report (Unpublished).
[2] Cochran, W. G. (1953). *Sampling Techniques*. Wiley, New York; Chapman an Hall, London.

* Since this article was written a colleague has used this approach to locate a sample adults 'with friends or relatives abroad'. A postal approach with two reminders achieve a response of 89% for a two-stage sample selected from the Electoral Register. Interviewi followed.

APPENDIX
The Questionnaires and Letters

(i) *The Postal Questionnaire—Front*

HEN AND PIG-KEEPING INQUIRY

If you have not kept any HENS, DUCKS, GEESE or TURKEYS at any time during the last 18 months, please write 'NONE' here and turn over .....

If you have, or have had, any poultry, *even if it is only one or two hens*, please answer questions 1 to 4.

1. How many fowls have you, or did you have,
   NOW? .....
   at the beginning of June this year (1954)? .....
   at the beginning of June *last* year (1953)? .....

2. Of the fowls you have NOW, how many are
   chicks under 6 months? .....
   cocks or cockerels? .....

3. Have you NOW got any ducks, geese or turkeys?
   Number of ducks .....
   Number of geese .....
   Number of turkeys.....

4. If you had any hens at the beginning of June 1953 (while feeding stuffs were still rationed) please show by writing 'Yes' against one of these answers whether you—

Managed without the feeding-stuffs ration for your hens then .....

Got your ration through the County Agricultural Executive Committee .....

Got your ration through the Food Office.....

PLEASE TURN OVER

To FARMERS and others completing Agricultural Returns

In a few cases we may have sent this form to someone who normally completes an Agricultural Return. If this should be so in your case, would you please merely write 'Yes' below and return this form to us? It is then unnecessary for you to answer the other questions.

Agricultural Return is completed .....

1

(ii) *The Postal Questionnaire—Back*

If you have not kept any PIGS at any time during the last 18 months, please write 'NONE' here .....

If you have kept any pigs please answer questions 5 to 9.

5. Are you now a member of a pig club? .....

6. How many pigs have you, or did you have,
   NOW? .....
   at the beginning of June this year (1954)? .....
   at the beginning of June last year (1953)? .....

7. Of the pigs you have NOW how many are
   Sows for breeding .....
   Sucking pigs .....
   Other pigs under five months .....
   Other pigs over five months .....

8. If you have given up keeping pigs since the beginning of June 1953 (last year) please say when you stopped, and why:
   .....
   .....
   .....

9. If you were keeping pigs at the *beginning of June* 1953 (while feeding stuffs were still rationed) please show by writing 'Yes' against one of these answers whether you
   Managed without the feeding stuffs ration for your pigs then .....
   Got the ration direct through your County Agricultural Executive Committee .....
   Got the ration for your pigs through a pig club .....

Are there any further comments you would like to make?
.....
.....

PLEASE RETURN THIS FORM AS SOON AS POSSIBLE IN THE STAMPED ADDRESSED ENVELOPE PROVIDED. PLEASE DO NOT DELAY EVEN IF YOU CANNOT ANSWER ALL THE QUESTIONS—WE WILL BE VERY GLAD OF ANY INFORMATION THAT YOU CAN GIVE US.

2

(iii) *Original Covering Letter*

16th November 1954

Dear Householder,

I am writing to ask for your help in an inquiry we are making on behalf of the Ministry of Agriculture about poultry and pig-keeping.

You will know that the Ministry gets full returns from farmers about poultry and pigs kept on farms. But a large number of people who are not farmers keep some hens or pigs and together they make an important contribution to national supplies. We are making this inquiry because the Ministry needs more information on just how big this contribution is now and also on the changes that have taken place since the rationing of feeding-stuffs ended.

Since there are no up-to-date records of pig- or poultry-keepers other than farmers, we realise that we will be writing to some people who have not kept any hens or pigs and indeed have no opportunity for doing so. If this applies to you, will you please still return the form putting 'none' at the top, as we would like to know this.

If you have any comments to add which you think will help to explain your answers to the questions, we would be very glad to have these. Any information that you give us will be treated as confidential and in no circumstances will names be mentioned in our report to the Ministry.

We shall be very grateful for your help.

(iv) *First Reminder Letter*

24th November 1954

Dear Householder,

A week ago I wrote to a number of householders, including yourself, asking them to say whether or not they had any poultry or pigs.

Most of the people have already replied but I have not yet heard from you. It would be a great help if you could spare a moment to complete the form now and let me have it back.

(v) *Second Reminder Letter*

7th December 1954

Dear Householder,

Hen and Pig Keeping Inquiry

About two weeks ago I wrote to you about an inquiry we are making for the Ministry of Agriculture. As I have not yet heard from you I expect the form never reached you or has got mislaid, so I am sending you another form with a stamped envelope. I hope you will be able to spare a moment to send it off.

NOTES AND COMMENTS

Readers interested in the methods of operational research, most of who will already be familar with publications on this subject in the Unite Kingdom and the United States, may like to know of the publication of new journal, *Revue de Recherche Opérationelle*. This is the official organ of th Société Française de Recherche Opérationelle, and its main objects ar giving information to members, the dissemination of ideas and methods, an the establishment of periodic contacts between its readers. It seeks to appea to a varied public, both executives and research workers. Although th methods of operational research are not exclusively statistical in character there will inevitably be many points of contact between the new journal an *Applied Statistics*, and we wish it every success. The publishers are Compagni Française d'Éditions, 40 Rue du Colisée, Paris 8, and the annual subscrip tion (four issues) is 2000 francs.

★ ★ ★

Karl Pearson was born in 1857. He is hardly known to post-war statist cians, since he retired from his chair at University College, London, in 19 and died in 1936, but to many of those who have passed middle age he gratefully remembered as The Professor. For the first quarter of this centu the development of statistical method, its application to the many depar ments of biology, and the beginning of its application to industry owed almo everything to K.P. The centenary of his birth was celebrated in a lectu given at University College, London, on 13th May 1957 by Professor J. B. S Haldane. Professor Haldane acknowledged the justification for some of th criticisms of Pearson's work, particularly in connection with theories heredity and evolution, but he remarked that in search of a *theory* he devis *methods* indispensable in any discussion of evolution and, indeed, in an serious application of statistics. Similarly, Columbus's theory of geograph was at fault; he set out for China and discovered America! Referring to th long-windedness of some of Pearson's algebra, Professor Haldane noted tha few ships now follow Columbus's course across the Atlantic, but we do no now regard him as a failure for these reasons. He referred to the many sidedness of Pearson's early work and career—his wranglership in mathe matics, his brief incursion into law, his writings on the history of art, hi writing of *The Grammar of Science*, and his radical political activities are onl some of the sides. In his lecture, which, we understand, will be published i *Biometrika*, Professor Haldane also gave some assessment of Pearson's work o heredity and biometry.

It is too soon to assess his work and influence definitively and in detail; bu he was a giant, and it is meet that his centenary should be celebrated.

MEETINGS OF SECTIONS OF THE ROYAL STATISTICAL SOCIETY

INDUSTRIAL APPLICATIONS SECTION

BIRMINGHAM AND DISTRICT GROUP

1 January 1957. A Flexible System of Sampling Inspection: G. A. arnard

February 1957. Statistical Bases of a Study of the Effects of Television: V. A. Belson

March 1957. The Application of Statistics in a Brewery: Miss S. V. unliffe

GLASGOW GROUP

1 February 1957. Measuring the Development of Reliability in Machines: . Morris

March 1957. Population Problems of a New Town: W. R. Matheson

LEICESTER GROUP

January 1957. The American Approach to Quality Control: C. S. Smith

1 February 1957. Electronics and the Automatic Office: J. C. Masters

1 March 1957. A Review of some New Techniques: C. H. Leigh-Dugmore

LONDON GROUP

February 1957. Experimental Design: M. J. Moroney

March 1957. Elementary Games Theory: S. Vajda

MERSEYSIDE GROUP

January 1957. Some Aspects of the Theory of Queues: F. Downton

February 1957. The Latest Developments in Quality Control in the SA: C. S. Smith

March 1957. A Statistical Approach to Stocktaking: E. S. Milner and Denman

NORTH-EASTERN GROUP

January 1957. Accidents and Road Layout: J. C. Tanner

February 1957. Testing Taste and Smell: W. A. Pridmore

March 1957. The Accuracy of Statistics calculated from Non-random ries, with Special Reference to Industrial Applications: G. H. Jowett

SHEFFIELD GROUP

February 1957. The Background to a Coal Sampling Specification: C. Tomlinson

March 1957. Accident Control Charts: D. W. Widginton

SOUTH WALES GROUP

14th January 1957. Monte Carlo Methods and Industrial Problems: W. Jessop

19th February 1957. Applications of Linear Programming: H. C. Macken

11th March 1957. Electronic Computers for Industrial Statistical Problem D. G. Owen

STUDY SECTION

LONDON GROUP

12th December 1956. Measuring the Effects of Advertising: D. H. Pickar

9th January 1957. Economic Prospects for 1957: J. Ryan (chairman), P Bareau, W. B. Reddaway, H. Rose

13th February 1957. The Statistical Aspects of Census Taking: B. Benjami

BRISTOL GROUP

13th December 1956. The TAM Service: P. Paine and A. Carden

Mr Carden described the methods used by Television Audience Measure ment Ltd. for collecting data on the size of ITA audiences in each trans mission area, and gave definitions of 'primary' and 'fringe' areas. H discussed the problems involved in devising the sampling schemes used fo selecting the individuals and households taking part in the firm's survey. Mr Paine then spoke of the ways in which the data are used for studying th television medium and for planning advertising schedules, giving illustratio of the methods by which the information is summarised and presented.

17th January 1957. Statistical Methods in Road Accident Research: I Garwood

Dr Garwood dealt with two methods of studying accident causatio firstly the collection of data on the causes of individual accidents, an secondly the use of information on the frequencies of accidents in differe circumstances. He showed graphs illustrating the effect of such factors driver's age and rainfall at different times of the year on accident rate, a described some of the work done by the laboratory in devising safety me sures, e.g. pedestrian crossings and improvements in the design of roa junctions.

21st February 1957. Some Uses of Economic Statistics in Industry: R. Savag

Mr Savage discussed the sphere of influence of the economist in indust and spoke of the use of statistical techniques in dealing with some of th economic problems confronting a firm. He gave as examples of such problem the assessment of the demand for the product with which a given industr deals; concern about supplies of raw materials; and wage negotiations, whe employers require information on the cost of living index and the reasons fo increases in prices.

MEDICAL SECTION

29th January 1957. Statistical Problems of the National Coal Board Medica Service: F. D. K. Liddell

26th February 1957. The Provision of Hospital Beds: D. J. Newell an A. T. Elder

26th March 1957. Arterial Blood Pressure in Population Samples: P. D. Oldha

BOOK REVIEWS

Allen Wallis and Harry V. Roberts: *Statistics: A New Approach*. Glencoe, Illinois: The ee Press, 1956. Pp. xxxviii + 646. $6.

'he claim made in the title of this book that it represents a new approach uite justified; this is something new in statistical textbooks. It is concerned nuch with the 'know-how' of statistical investigation and inference as with nniques.

'he first 344 pages, about one-half of the book, are arm-chair reading, n scarcely an equation and only a few simple arithmetical calculations. t I, on 'The Nature of Statistics', describes fields of application, common rs in the interpretation of statistical information, the recording and raisal of data (these two chapters are almost as much essays in scientific hod as in statistics), and the basic ideas of probability and sampling. t II is on 'Statistical Description' and introduces charts, frequency ributions, averages, and common measures of variability and association. t III is on 'Statistical Inference' and in the first chapter discusses ran- nness and probability. So far, the reader has become aware of only the t elementary of statistical techniques, but has had all the problems that e in drawing conclusions from data discussed in an easy, leisurely, and erting way, with a wealth of illustrations. The following two quotations e to illustrate the tone of the discussion—many others could be given.

'Inexperienced users of statistics, for example, will often ask questions ke, "Should I compute the arithmetic mean or the median of my data?" 'he prior question, as we have tried to emphasize, is "What am I trying o find out?" Only after that question is clearly answered can specific echniques be chosen. Thus, the answer may be that *neither* the arithmetic ean *nor* the median is appropriate, but that something else—for example, ne entire frequency distribution—is required. Particular statistical ethods are simply tools, the skilled use of which requires insight into the ubject matter and a clear conception of the problem.'

'The reading of statistical studies almost never demands extensive com- utations with elaborate equipment. Another kind of calculation—a mental approximation"—is much more useful. It is unfortunately true at many producers of statistical studies are careless about calculations, nd that even with great care, as we have seen, mistakes sometimes occur. ence, for serious reading one ought to form the habit of checking cal- ulations roughly, when this is possible, to see if gross errors have been made.'

he authors have scoured far and wide for examples: newspapers, journals, erto unrecorded experiences, and, of course, the Kinsey Report. They a passage in Thucydides to illustrate the mode of a frequency distribution, the following example is not the only striking one in the book.

xample 141A. Pearl Harbor

An erroneous rejection of extreme observations profoundly affected the urse of history when, shortly after 7 o'clock on the morning of December 1941, the officer in charge of a Hawaiian radar station ignored data lely because they seemed so incredible.'

After page 344 the book continues in much the same key as before exc that statistical techniques and their uses are explained, and the pace is fas The subjects dealt with are sampling distributions, the normal distributi decision procedures, tests of significance, estimation, the design of experime and surveys, statistical quality control, correlation and regression in th various aspects, time series, and short-cut procedures. A lot of material packed into this half of the book, and much of it requires to be studied rat than read.

This book is the mature product of years of experience in the teaching statistics to students of all kinds. It *teaches* statistics and provides a su stantial contribution to a liberal education. It is not statistics for economis or biologists, or industrialists, or any other 'ists'; it is just statistics. Teache will find it bristling with ideas on presentation and with material with whi to illustrate their lectures; students will find it stimulating as well as unde standable; readers interested in the processes of gaining knowledge of th universe, especially the social part of it, will find it interesting and helpfu A wide circulation for this book will do much for the cause of statistics an all that goes with it.

L. H. C. TIPPET

A. R. Ilersic: *Statistics and their Application to Commerce, 11th Edition.* London, H.F.L. (Pu lishers) Ltd., 1956. Pp. xii + 401. 25s.

With the appearance of a new edition of this textbook the author h taken the opportunity to make substantial additions to those chapters deali with sampling methods and survey problems and to bring up to date t section covering economic statistics (including index numbers). The chapte on methods have been modified only slightly except for that dealing wi correlation, which has been extensively redrafted. In general, it remains elementary textbook designed primarily for 'those students working professional examinations in which statistics in its various guises is set as subsidiary subject'. It is essential that its value should be judged in rega to this aim; when viewed in this light the book does provide a sufficie coverage both of those methods required for such examinations and economic statistics and survey procedures.

The methods are in most cases described in a predominantly verb fashion with only a limited resort to algebraic notation. Bearing in mind t class of reader for whom the book is intended this is understandable, but does on occasions make for over-elaboration and does not always facilita a clear exposition of some concepts—as, for example, the explanation of th short-cut methods of calculating the arithmetic mean and standard deviatio The use of symbols to supplement (if not to supplant) such explanatio would, I feel, assist many readers, even those of a non-mathematical min As the author says, such symbols are 'abbreviations of simple procedur which would be cumbersome if expressed in simple English', but he himse has not always been able to avoid such cumbersomeness. It is unfortunat that in the text there are instances of lack of consistency where notation ha been introduced. Examples occur in Table 30 on p. 131, dealing with th calculation of the standard deviation, where the headings constitute mixture of symbols and verbal descriptions, and again in the bivariat frequency table on p. 239, where, in those additional columns used fo calculation, the transformed variable (in the form $[x - A]/u$) has not bee

ficiently distinguished from the variable in its original form. This Table o contains a more serious error in notation arising from confusion between e two variables x and y. Those columns covering the calculations of $\bar{x}$ and have been labelled in terms of y, and similarly those covering the mean d standard deviation of y have been labelled in terms of x.

The chapter covering correlation and regression has been rewritten and nplified since the previous edition. Incidentally, why was not this chapter roduced earlier in the book, and perhaps inserted after that on the normal rve? The section on the significance of r could then have been transferred the subsequent chapter, which deals with significance tests generally. The cussion of regression has been reduced and brought more into line with e book's general level, the fairly lengthy section on the standard error of e estimate having been omitted. The calculation of r in terms of this tistic has been left out and merely stated in its more conventional form. eel that the introductory descriptive remarks on the meaning of correla- n are still not sufficiently clear in their distinction between statistical rrelation, in the sense of an association of variables, and meaningful rrelation. The author states, for example, 'that the mere fact that two riables appear to move in sympathy does not establish conclusively the stence of correlation'—this may be true in an interpretative sense, but is ely not so in a purely statistical sense. Fortunately, Ilersic has not for- tten to emphasise the dangers of the indiscriminate use of the correlation efficient.

When considering the statistical sources which are covered by this text- ok it has to be realised that few of the students for whom it is intended ve ready access to the wide range of published data on index numbers d economic statistics (as found in Government Journals and other publi- ions and in learned journals). For such students a book on statistical thods which includes a survey of sources is invaluable, particularly when is remembered that the few textbooks in this field are either too com- ehensive or unsuitable so far as coverage is concerned. Ilersic has described sources which will be found of most use to examination students. These lude man-power, national income, production, and population in the apter entitled Statistical Sources; and among Index Numbers those ating to prices, wage rates, production, and retail sales. To the latter apter have been added, in this edition, those index numbers covering rling-area trade. In addition to the sources referred to above the author given a few useful references for more detailed reading. Although this e of information must of necessity suffer from the danger of becoming dated, it does provide the student with a basic knowledge generally ficient for examination purposes. Unfortunately, it is unlikely to be equate in the field of practical application.

Time Series have been dealt with in a chapter which is inserted between se on Index Numbers and Statistical Sources. This might have been re appropriately placed immediately before the former.

The chapters dealing with sampling methods and survey problems have n rewritten to incorporate the material which has been published since previous edition, but remain broadly the same as before. As with nomic statistics, this section provides a useful review of existing methods, l supplies a limited number of references to the more prominent papers l publications in this field.

S. BURKE

Robert E. Johnson and Doris N. Morris: *Guide to Elementary Statistical Formulas*. New Yo and London: McGraw-Hill Book Co. Inc., 1956. Pp. 102. 22s. 6d.

This book is an attempt to bring together the elementary mathematic and statistical formulae which are frequently encountered in business. is not a textbook, but it includes a selection of references for further stud The fact that almost half the available space is devoted to index numbe and time series gives some indication of the public for which it is primari intended. There is an edge index to facilitate reference to the sectio required. D. G. BEEC.

PUBLICATIONS RECEIVED

Archivio Economico dell'Unificazione Italiana, *Volumes* 1–4. Turin: Industria Libraria Tipo grafica Editrice, 1956.

Yearbook of International Trade Statistics, 1955. New York: Statistical Office of the Unite Nations, Department of Economic and Social Affairs, 1956. Pp. 720. 50s.

Mental Health Research Unit, Technical Report, 1955. New York: State Department of Menta Hygiene. Pp. 127.

Calvin F. Schmid and Vincent A. Miller: *Enrollment Forecasts, State of Washington:* 1957–1965 Seattle: Washington State Census Board, 1957. Pp. viii + 78.

THE SAMPLING OF ROAD TRAFFIC

JOHN C. TANNER

Road Research Laboratory,
Department of Scientific and Industrial Research

The study of road traffic may involve sampling problems; in this article Mr Tanner shows how adaptations of fairly simple methods may be used to obtain representative samples.

duction

his paper describes three sample investigations in which the 'popu-n' being sampled was the road system of Great Britain, the object g to provide information about the roads themselves or about the ic using them. The first was a sample count of junctions on the of area; the second investigation comprised a survey of the ibution of traffic; and the third investigation was carried out to sure trends in road usage. Although the methods of sampling used fairly simple ones, their application has presented a number of esting problems, and the emphasis of the paper is on these prob- rather than on the results obtained.

Roads of Great Britain

or administrative purposes, roads in Great Britain are divided into main classes, and the mileages of each class in towns and in rural s are given in Table I.

TABLE I

Mileage of road, Great Britain

Class of Road	Towns	Rural Areas	Total
ınk	1 600	6 700	8 300
ss I	6 100	13 500	19 600
ss II	3 800	13 900	17 700
ss III	5 300	43 500	48 800
classified	37 100	56 600	93 700
Total	53 900	134 200	188 100

n these 188 000 miles of road, about 6 million motor vehicles are en a total of approximately 50 000 million miles per year; in other ds, 50 000 million motor-vehicle-miles are performed each year. ill be seen later how this traffic is distributed among the various s.

Sampling Methods Available

Two main methods have been used to select a sample of the roa system. In the first method the roads are regarded as a one-dime sional array with a total length of 188 000 miles. If one imagines a the roads laid out end to end, random numbers can be used to sele a random sample of any desired size. This might be a sample of poin on the road system, or it might be a sample of, say, one-mile stretche depending on requirements. This method presupposes the availabilit of sufficiently detailed lists of roads and their lengths.

The second approach is to regard the road system as a two-dimen sional array, i.e. as laid out on a map. The area of the country ma be divided into a large number of small squares and a sample take by considering all roads lying within a random sample of these squares. This method appears less suitable for selecting points as distinct from lengths of road, but it might be possible to use it for this purpose. Some of the advantages and disadvantages of these two approaches t sampling will be apparent from the investigations to be described.

Sampling on the Basis of Area

The first investigation to be described was essentially an explorato one, but it serves to illustrate the method. The problem was to esti mate the number of road junctions in the country as a whole, dividin them up by type. For simplicity, however, it was decided to limit th scope of the work to junctions of A and B roads in England and Wal (normally A roads are trunk or class I roads, and B roads are class I roads). The area method was used because all the information require was contained on easily available maps.

The National Grid used by the Ordnance Survey superposes on th map of Great Britain a rectangular grid of 100-km squares, each squar of which is subdivided into one hundred 10-km squares. England an Wales are completely covered by twenty-eight of the 100-km square but a number of them consist mainly of sea, and by superposing number of these, the effective number is reduced to twenty.

A five per cent sample was drawn, with five 10-km squares chose at random from each of the 100-km squares. There were thus 10 sample squares. The appropriateness of this procedure is considere later. The number of junctions in each of the sample squares was then counted from maps. The frequency distribution of the number o junctions in the 100 sample squares was as shown on the following page.

Altogether there were 651 junctions, and since the sampling fraction was one in twenty, the estimate of the total number of junctions o A and B roads in England and Wales is $20 \times 651 = 13\,020$. However the number of junctions varied widely from one sample square to another: this results in a rather high coefficient of variation, abou 20 per cent, for the estimate 13 020.

Although the number of junctions varies widely from one sample square to another, the proportions of various types vary much less. Fo

Number of Junctions per 10-km Square	*Frequency*
0	24
1	17
2	10
3	6
4	5
5	8
6	7
7	7
8	2
9	2
10–19	5
20–29	3
30–39	1
40–49	2
149	1
	100

ıple, no significant variations were found between the proportions ree-way junctions in the various squares.

ne results of this study suggest several ways in which the methods might have been improved, apart from the obvious one of increas- he sampling fraction. In the first place, more precise results would been obtained by using a larger number of smaller sample res, say 10 000 1-km squares instead of 100 10-km ones.

nother refinement would have been to use a covariate, such as th of road, in the analysis. If the number of junctions were found e linearly related to the length of road in the sample squares, one d use the known total mileage of road in the country to correct number of junctions in the sample for the mileage of road in the ole squares.

stly, the sample might profitably have been weighted by varying umber of sample squares from one 100-km square to another, with e of them in the 100-km squares containing greater road mileages.

rvey of the Distribution of Traffic

fore this survey is described, the terms used will be defined:

The *flow* at any point on a road is the number of vehicles passing at point per unit time.

The *vehicle-mileage* on any road or set of roads is the number of les driven by vehicles on those roads in a specified period.

e draws a very large random sample of points on a set of roads measures the flow at each point, one can calculate an average flow plot a distribution of flow for the set of roads. The vehicle-mileage

defined above is this average flow multiplied by the total length of t set of roads in miles.

The main objects of the survey were to study the average flows a vehicle-mileages on various sets of roads and to study the distributi of flow over the road system as a whole.

Sampling Method. A two-stage sampling method was used. In t first place four local authority areas were selected, and then a stratifi random sample was drawn from each area. Since the four areas we chosen on an arbitrary basis, and since in any case four areas are i sufficient to give results representative of Great Britain, it is necessar at least until a late stage in the analysis, to treat the survey as fou separate investigations.

The four areas selected were Buckinghamshire, Dorset, Lancashir (excluding county boroughs), and Leicester County Borough. Thes areas cover a wide range of conditions and, taken together, the rati of town to country is about the same as the ratio for the whole of Grea Britain.

In each of these areas fifty census points were selected. It was firs decided how many of them should lie on each of ten strata consistin of trunk, class I, class II, class III, and unclassified roads, in built-u and in non-built-up areas (defined by the presence or absence of speed limit). This was done, so far as possible, to give the greate possible precision for the estimated distribution of flow over the roa as a whole. The number of points in each stratum having been fixe lists of roads and their lengths were used in conjunction with tabl of random numbers to allocate the points at random within strat The method employed was to allocate one point in each stratum random, and place the remainder at equal intervals up and down t list of roads. Since it is most unlikely that there is any periodicity i the flow as one goes down the lists of roads, this procedure shou provide a sample with somewhat better geographical coverage than truly random one, without any compensating disadvantage.

The next and main stage in the investigation was to estimate th traffic flow at each of the fifty points in each area. Briefly, the meth involved counting for two 1-hour periods or for two 2-hour perio during one week (different in different areas) at most of the point scaling up the results to the Monday to Friday average for that wee by means of continuous counts for the whole week at one point, an scaling the figures for the week to the average figure for weekdays i the ensuing twelve months with the help of quarterly counts at fou or five points. The figures were finally corrected to the 1954 level b means of trends assessed independently. No precise measure of accurac of the final result can be calculated, but errors in the scaling processe will not seriously affect the distribution of traffic between differen roads, although they will affect the general level of the figures. Al subsequent analysis is carried out on these figures for an average week day in 1954. Some typical results of the survey are as follows.

ıe average flows on the ten groups of road, averaged over the areas (weighted according to road mileage), are given in Table II.

TABLE II

Number of vehicles (including pedal cycles) per weekday

Class of Road	Built-up Areas	Non-built-up Areas
ınk	6 100	6 100
ss I	4 200	2 400
ss II	2 200	900
ss III	1 500	400
classified	900	200

figures for trunk roads are based on relatively few census points, ıat the agreement between the flows on these roads in built-up in non-built-up areas is no doubt fortuitous. Although flows are ı higher on the more important roads, the mileage of such roads s only a small proportion of the total. By multiplying the flows oad mileages and reducing to percentages, the figures given in e III are obtained.

TABLE III

Percentage of total vehicle-mileage

Class of Road	Built-up Areas	Non-built-up Areas
ınk	7	16
ss I	16	12
ss II	6	4
ss III	6	8
classified	20	5
Total	55	45

milar results can be calculated for each type of vehicle. Table vas derived from the results for pedal cycles.

TABLE IV

Flow of pedal cycles as percentage of flow of motor vehicles

Class of Road	Built-up Areas	Non-built-up Areas
ınk	13	4
ss I	19	18
ss II	37	18
ss III	54	28
classified	76	64

stribution of Traffic. For each class of road in each area, the mileage ad carrying less than any given flow can be estimated from the sponding sample proportion. For an area as a whole, these

mileages for individual strata can be added together and the res expressed as a percentage of the total mileage. Fig. 1 was obtain in this way; the flow is plotted on a logarithmic scale and the pe centages on a normal probability scale. When the results are so plotte

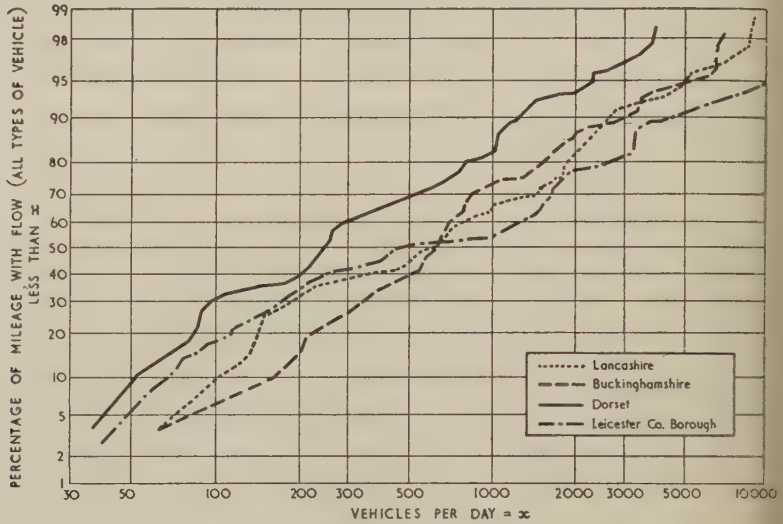


FIG. 1. Cumulative frequency distribution of flow.

a straight line implies that the logarithm of the flow is distribute according to the normal law over the population of road mileage. will be seen from Fig. 1 that for each area the distribution gives convincing straight line over at least two cycles of the graph, except f erratic variations attributable to the small number of census point Further, with the possible exception of Leicester County Borough, eac of the lines has approximately the same slope; this means that th standard deviations of the distributions of log-flow are the same.

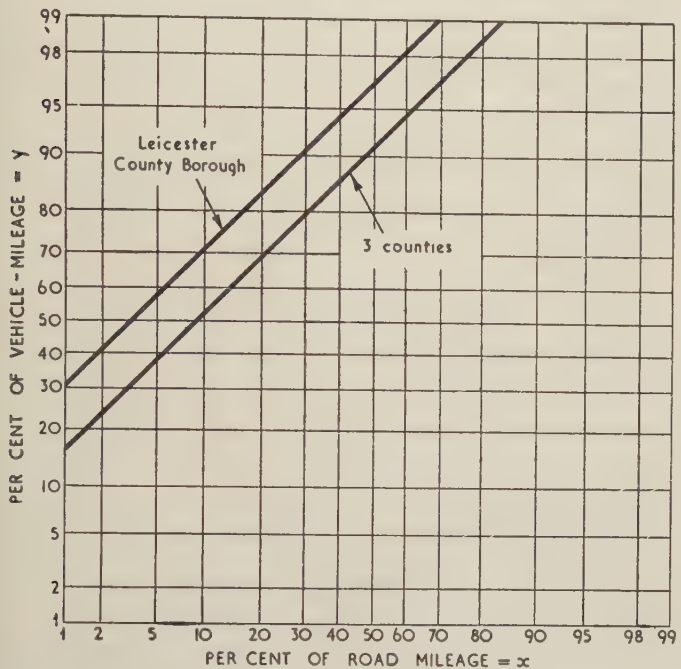
Another type of relationship shows the percentage y of the vehicl mileage carried by the busiest x per cent of the road mileage. It happens that with the lognormal distribution of flow already demo strated, the relationship between y and x is a straight line if both ar plotted on normal probability scales, as shown in Fig. 2. This diagram is very closely related to the Lorenz curves used by economists t measure income or business concentration.

Measurement of Trends in Road Usage

The third investigation to be described has as its object the measure ment of changes from day to day and from year to year in the vehicle-mileage in the country as a whole. This requires continuous (o intermittent) counts at a number of places, and such counts wer started at the beginning of 1956. It is thus too early to give any results

It was decided for convenience to make counts by means of abou

permanently sited automatic recorders. A recorder consists of a er tube stretched across the carriageway and a counter mechanism e roadside. Each axle of each passing vehicle sends a puff of air the tube, and this operates the counter. The instruments are not



Percentage of vehicle-mileage, y, carried by the busiest x per cent of road mileage.

ly sensitive to pedal cycles. To obtain daily totals, each site is d early each day. The problems to be considered thus fall into roups: (1) the location of the census points, and (2) the analysis records.

e general idea behind the design of the census is that it should ossible to estimate the vehicle-mileage (strictly, half the axle-ge of motor vehicles) in the country on each day. Changes from eriod to another can then be calculated without difficulty. This sts that a stratified random sample of points on the road system uired, similar to the samples for the four areas in the survey of stribution of traffic already described. It was decided, however, clude unclassified roads; the classified roads carry about 80 per f the motor-vehicle-mileage and their length is only about 50 per f the total.

e first stage was to fix the number of points to be allocated to the strata: trunk, class I, class II, and class III roads, in urban and al areas. This was done in such a way as to minimise the standard of the estimates of trend in vehicle-mileage over the roads as a . Although a valid sample of points would have been obtained

by placing the points at random within strata, it was thought desirabl for statistical and other reasons, to balance the sample, so that ov the fifty points as a whole a reasonable number of points should lie i each of the ten Ministry of Transport Divisions, and so that withi each Division the points should be reasonably distributed between th classes of road, and between built-up areas and non-built-up area This presents both practical and theoretical difficulties. An explan tion will first be given of what was done and then reasons will be give for believing the method used to be acceptable.

Table V shows the 'expected' number of points in each stratum i each Division, subject to the number of points per stratum alread determined. Within each stratum the expected numbers are propo tional to the road mileages in each Division.

TABLE V

Expected number of points in each stratum in each Division

Division	Urban Roads					Rural Roads				
	Trunk	Class I	Class II	Class III	Total	Trunk	Class I	Class II	Class III	Total
A	0·23	0·61	0·18	0·24	1·26	1·94	2·87	1·61	1·63	8·05
B	0·35	0·48	0·24	0·24	1·31	0·84	0·66	0·64	1·50	3·64
C	0·21	0·52	0·23	0·32	1·28	0·76	1·14	0·70	2·23	4·83
D	0·25	0·58	0·25	0·28	1·36	0·52	0·99	0·54	1·14	3·19
E	0·46	1·29	0·48	0·50	2·73	0·13	0·21	0·14	0·28	0·76
F	0·21	0·28	0·22	0·18	0·89	0·45	0·67	0·63	1·17	2·92
G	0·33	0·41	0·24	0·23	1·21	0·67	0·78	0·50	1·20	3·15
H	0·31	1·24	0·42	0·42	2·39	0·64	0·64	0·53	1·12	2·93
I	0·35	1·00	0·42	0·36	2·13	0·53	0·39	0·28	0·73	1·93
J	0·30	0·60	0·31	0·24	1·45	0·51	0·65	0·42	1·00	2·58
Total ..	3	7	3	3	16	7	9	6	12	34

It was decided that the total number of points in each Divisio should be the integer nearest the expected number (9·31, 4·95, and s on); Division E, however, was given 4 instead of 3 to give the righ total, 50. The points allocated to each Division were then divide between its urban and rural parts. This fixed the marginal totals i the urban and rural parts of the above table, and it remained t allocate the points to cells.

This was done one point at a time as follows. First, in those cell where the expected number is greater than unity, one or more point were allocated to it and the marginal totals were reduced accordingly This left all the cells with expectations between 0 and 1, with integra column totals adding up to 13 in urban areas and 20 in rural areas but with row totals different from those already fixed. Perhaps th most satisfactory procedure would have been to make random alloca tions within each column until a set of allocations was found whic gave the correct row totals. This, however, would have taken a ver

; time with tables as large as these. Instead, the row or column in urban block with the greatest total (class I with 4) was allocated oint, the cell being selected at random with a probability for each oortional to the (reduced) expectation. The corresponding mar- l totals were reduced by unity and the procedure was repeated the row or column with the next highest total. By this means all points were allocated in the urban areas, and a similar procedure followed for rural areas. To improve the distribution by class of l in each Division a point was moved from E, urban, trunk, to E, an, class III; and one from J, urban, class III, to J, urban, trunk. Within each cell, points were then allocated by means of random nbers and lists of roads. In some Divisions, where too many points pened to occur in a particular county, one (or more) of them was llocated to one of the other counties in the Division.

he above procedure, though not strictly justifiable, was thought to cceptable for the following reasons. In the first place, once Divi- and stratum have been fixed, the selection of each point follows odox lines, apart from one or two unimportant modifications: it is first part of the procedure that may be questioned. The allocation oints to Divisions, however, is as satisfactory as most, if not all, odox random allocations. Similarly, within any one Division, the cation between the various strata is as good as most random alloca- s. The sort of criticism which might be made is that the method s too few points to the cells with small expectations, and that these l to be minor roads in towns in sparsely populated Divisions, where ds may be exceptional. However, any such bias in the allocation oints must be small because of the check that was made of the cation of points to strata within Divisions, and in any case such is unlikely to be highly correlated with differential trends on erent sets of road.

he counters at present in use may suffer from various internal s, and may be affected by snow, frost, parked vehicles, wilful age, and so forth. This presents some difficult problems, not only ne estimation of missing values, but also in the assessment of the worthiness of recorded numbers of vehicles.

aking the latter point first, some of the daily flows recorded can ejected immediately; in many cases, however, it is not at all certain ther or not the figure is correct. The basic rule being followed is any weekday flow differing from the average for the month by e than 20 per cent is rejected; greater tolerances, 40 and 60 per , are allowed for Saturdays and Sundays. However, so many ptions have to be made to this rule that it may finally be decided bandon it. When it has been decided which daily totals to reject, the pted ones are entered on Cope-Chat cards, one per point per month. wo methods are being used to estimate the missing figures. The is to use the average of the figures for the same day of the week in nbouring weeks at the same point. The second is similar, but

results from other points are used to correct for real differences in the level of traffic between the weeks. The former method is quicker to use, but it tends to iron out real differences between days, due to factors such as weather. The second method is only satisfactory if there is a fairly high correlation between the day-to-day fluctuations at the defective point and the average of the other points. It is hoped to study the size of these correlations shortly. At present the simpler method of estimation is used for short breaks in the sequence of daily values, and the second method for long breaks. Precise, but somewhat lengthy, rules have been laid down to determine the number of alternative points and the number of neighbouring weeks to be used.

When all the missing values have been estimated, the average flows in the various strata will be multiplied by the length of road in them to give vehicle-mileages, and all further calculations will be carried out on these values. The standard errors of the estimates of vehicle-mileage and of changes of vehicle-mileage will be assessed in the usual way from the variations between census points. These standard errors will, it is believed, overestimate the true errors, but by an unknown amount. This is because the geographical distribution of points has been made better than random, whereas the standard errors are based on the assumption of randomness.

From the estimates of vehicle-mileage for each day, indices will be calculated month by month and year by year to show trends in road usage on weekdays, Saturdays, and Sundays, and over the months as a whole. Provided that the results are sufficiently consistent, separate indices will be calculated for each part of the country, for each class of road, and for built-up and non-built-up areas.

Acknowledgements

The Laboratory is grateful to the highway authorities concerned and to the Ministry of Transport and Civil Aviation, without whose co-operation the traffic surveys described could not have been undertaken. The work was carried out as part of the programme of the Road Research Board of the Department of Scientific and Industrial Research. The paper is published by permission of the Director of Road Research.

NORMALITY IN QUALITY CONTROL CHARTS

PETER G. MOORE
University College, London

The assumption of normality in the interpretation of quality control charts may be accepted uncritically by some and viewed with scepticism by others. In this article Dr Moore shows how certain departures from normality can affect the probabilities associated with control limits calculated by normal theory. These examples should serve as a warning against the too rigid interpretation of such control limits.

duction

he setting up of a quality control chart for the means and ranges nall samples is nearly always based on the mean and range in a ious series of small samples from the product that is being manu- ured. It is also customary to assume that the statistical distribution ie particular characteristic with which we are concerned is normal he sense of Gaussian) or very nearly so. Using the assumption of nality it is possible to make statements to the effect that the propor- of articles with the characteristic concerned falling outside the age of the whole bulk of articles by more than some multiple of standard deviation is approximately constant. Sometimes, how- , it is not emphasised that if the basic distribution is not normal e error will be introduced in the charts by using the normal-curve ors provided in the quality control literature.

y no means all the distributions that arise in situations where ity control techniques are commonly applied could be considered e normal. To measure the degree of departure from normality of tribution two quantities are usually calculated from the moments, f the distribution. These are

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} \qquad \text{and} \qquad \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$$

quantity γ_1 measures the skewness of a distribution or its departure symmetry, and γ_2 measures the kurtosis or degree of 'peakedness' e distribution. In nine distributions, quoted by E. S. Pearson[1] and lied by W. A. Shewhart, of observations collected by the Bell Tele- ie Laboratories in connection with their work, eight show some of skewness ranging up to a value of γ_1 nearly equal to one. In e cases there is also considerable peakedness, shown by γ_2 being ter than one. Student[2] comments: 'many if not most routine yses have a leptokurtic (i.e. γ_2 greater than zero) form of error

system'. If quality control charts are set up for samples drawn from such non-normal populations using normal theory the various probability levels used will not be equal to the true probability levels.

To set up a control chart the mean and standard deviation of the population of items being manufactured are usually estimated from a series of samples already drawn. If the mean and range of each sample have been found, the average values of these two quantities may be used to provide estimates of the population mean and population standard deviation. Such estimates will be liable to two forms of errors. The first form of error will be due to sampling fluctuations. This form of error will always be present but will diminish as the number of observations used for estimation increases. For example, the coefficient of variation of the mean range in 25 samples of size 5 from a normal population is 7·4%, in 100 samples of 5 it is 3·7%, and in 500 samples of 5 it is 1·7%. It is clear that these errors diminish as the number of samples available increases. The second form of error is due to the parent population from which the observations are selected not being normal although normal theory is being assumed throughout. Errors due to this source are always present and are not eliminated when more samples are taken and may thus produce biased estimates of the population standard deviation. Such errors will be termed systematic errors.

In this paper we shall examine systematic errors for a number of non-normal parent populations, by considering first the effects of such distributions on the probability levels of the mean and secondly the effect on the estimation of the standard deviation using the range in the samples. Thirdly, an attempt will be made to combine these two effects by considering the overall effect of using normal theory both to estimate the population constants and then to set up the required limits for the control chart. The presence of any random sampling errors may reinforce or oppose these systematic errors, but it should be borne in mind that such random errors can always be reduced by taking further samples whereas systematic errors cannot be reduced in this manner.

Sensitivity of the Mean

The χ^2 distribution with 4 degrees of freedom is markedly skew and peaked, having a γ_1 of 1·41 and a γ_2 of 3. If random samples of size five are drawn from such a population the means of the samples will themselves be distributed with a mean of 4 and a standard deviation of 1·264 9. Now under normal theory we assume that the means are in fact normally distributed with this mean of 4 and standard deviation of 1·264 9. Hence a value, F_α, can be found that will cut off a proportion α of the distribution. However, in this case the exact distribution of the mean is known, being another χ^2 distribution but with 20 degrees of freedom and a scale factor. From this distribution the exact or true probability cut off at the point F_α can be calculated and compared with α cut off

he normal approximation. The results are given in Table I, where rresponds to the proportion at one end of the distribution; the final

TABLE I

Nominal and true probabilities

α	True Probabilities		Two Tails	Error
	Upper Tail	Lower Tail		
0·10	0·107 1	0·080 5	0·187 6	−0·012 4
0·05	0·063 9	0·025 1	0·089 0	−0·011 0
0·025	0·039 5	0·006 0	0·045 5	−0·004 5
0·01	0·021 8	0·000 4	0·022 2	+0·002 2
0·005	0·014 3	0·000 0	0·014 3	+0·004 3

columns show the combined probability, based on two tails, com-ed with a nominal probability of 2α. From the table it is apparent t the upper tail will always have a probability greater than the ıinal, whereas the lower tail will have a probability lower than the ıinal level. If, however, the test is to be a two-tailed test then the ıbined probabilities are, with one exception, appreciably nearer to nominal level than either of the tail probabilities separately. The ticular distribution studied here is very noticeably non-normal and next example uses a χ^2 distribution with 8 degrees of freedom, for ch $\gamma_1 = 1$ and $\gamma_2 = 1{\cdot}5$, so that it is nearer normality and resembles ıe of the distributions quoted by Pearson. In this case two sample s, five and eight, have been considered and a procedure similar to t used above followed throughout. The results are given in Table and comparing first with Table I we notice that the probabilities a sample size of five are nearer to the nominal values in this case

TABLE II

Sampling from skew population

α	$n = 5$			$n = 8$		
	True Probabilities			True Probabilities		
	Upper Tail	Lower Tail	Two Tails	Upper Tail	Lower Tail	Two Tails
10	0·106 7	0·088 4	0·195 1	0·105 5	0·092 0	0·197 5
05	0·061 3	0·034 1	0·095 4	0·059 2	0·038 4	0·097 6
025	0·036 2	0·011 7	0·047 9	0·033 9	0·014 6	0·048 5
01	0·018 9	0·002 3	0·021 2	0·016 9	0·003 6	0·020 5
005	0·011 8	0·000 5	0·012 3	0·010 3	0·001 1	0·011 4

than previously. This is not surprising, as it is known that the near the parent population is to normality the more nearly will the mea of samples of any fixed size be normally distributed. However, the are still considerable differences in the upper and lower tails by them selves, and once again it is only when the combined tails are considere that the nominal and true levels nearly coincide. A further improv ment is effected by increasing the sample size. Once again the ind vidual tail probabilities get closer to their nominal values, and th combined tails also show some improvement.

To recapitulate, these results suggest that if two tails are being co sidered the errors should not be large but that there may be big erro in considering either tail by itself. These errors depend on the degre of non-normality and also on the sample size, and can in general b reduced by using a larger sample size.

Effect of Variability Errors

When a quality control scheme is drawn up the estimate of variatio is usually obtained from a number of samples of the product that being examined. Suppose that, whatever the method that is used t estimate the standard deviation, an error is made. The effect of th error can be measured for differing circumstances; the example give here assumes that samples are drawn from a normal population an that control limits are set up using the correct mean but a standar deviation that is in error. Table III shows the results in the form

TABLE III

Effective probabilities with errors in standard deviation

Percentage Error in Standard Deviation	Nominal Probabilities		
	0·05	0·025	0·005
+10	0·035 2	0·014 5	0·002 3
+ 5	0·042 1	0·019 8	0·003 4
+ 2	0·046 7	0·022 8	0·004 3
− 2	0·053 5	0·027 4	0·005 8
− 5	0·059 1	0·031 3	0·007 2
−10	0·069 4	0·038 9	0·010 2

probabilities of means falling in the upper tail as compared with th nominal probability, using six different levels of error in the standar deviation. The errors would be exactly the same for the lower ta and are also independent of the size of sample that is being considered This implies that the error is not reduced by using larger-sized sampl and that any error introduced by an incorrect standard deviation wi

ys be present. The magnitude of the probability deviation is quite nd varies enormously according to both the nominal level involved the amount of error that has occurred in the standard deviation. with errors as small as 5% in the standard deviation it is possible ve a difference of 44% in the significance level involved. Whether s of practical importance depends, however, on the emphasis and pretation that is placed upon sample results falling outside desig- d limits. It serves, perhaps, as a warning that significance limits ld not be interpreted too rigidly.

ation of Standard Deviation

ne most common method of estimating the standard deviation in population for quality control purposes is to use the range in the ple of size n, i.e.

= range = highest observation — lowest observation in sample

then multiply the mean range from a fairly large number of ples by a factor, $1/d_n$, to convert it to an estimate of standard ation. The factors d_n are obtained by considering the distribution nge in samples from a normal population and are chosen so that

$$\mathrm{E}\left(\frac{w_n}{d_n}\right) = \sigma$$

re E stands for expectation or mean value in repeated sampling σ is the standard deviation in the population that is being sampled. ce if we consider the quantity $w_n/d_n\sigma$ this will have an expectation, mpling from normal populations, of unity.

we now draw samples from various non-normal populations but he normal factor d_n to estimate the standard deviation, an error be committed, and by finding $\mathrm{E}(w_n/d_n\sigma)$ under these circum- ces the magnitude of this error can be gauged. Five populations be used and sample sizes varying from two to ten will be considered. five populations considered, together with their appropriate mea- s of departure from normality, are

Exponential	$\mathrm{p}(x) = \mathrm{e}^{-x}$	$0 < x < \infty$	$\gamma_1 = 2$	$\gamma_2 = 6$
$^2{}_6$	$\mathrm{p}(x) = \frac{1}{2}x^2\,\mathrm{e}^{-x}$	$0 < x < \infty$	$\gamma_1 = 1{\cdot}12$	$\gamma_2 = 2$
$^2{}_8$	$\mathrm{p}(x) = \frac{1}{6}x^3\,\mathrm{e}^{-x}$	$0 < x < \infty$	$\gamma_1 = 1$	$\gamma_2 = 1{\cdot}5$
Double exponential	$\mathrm{p}(x) = \frac{1}{2}\,\mathrm{e}^{-\lvert x\rvert}$	$-\infty < x < \infty$	$\gamma_1 = 0$	$\gamma_2 = 3$
Logistic	$\mathrm{p}(x) = \dfrac{\mathrm{e}^{-x}}{(1+\mathrm{e}^{-x})^2}$	$-\infty < x < \infty$	$\gamma_1 = 0$	$\gamma_2 = 1{\cdot}2$

values of $\mathrm{E}(w_n/d_n\sigma)$ for these five populations are given in Table Some of the values under (*a*) and (*e*) have been previously given

TABLE IV

Values of $E(w_n/d_n\sigma)$

Parent Population	Sample Size (n)								
	2	3	4	5	6	7	8	9	10
(*a*)	0·886	0·886	0·891	0·896	0·901	0·906	0·911	0·916	0·919
(*b*)	0·959	0·959	0·961	0·963	0·964	0·966	0·967	0·968	0·969
(*c*)	0·969	0·969	0·970	0·971	0·972	0·973	0·975	0·975	0·976
(*d*)	0·940	0·940	0·952	0·966	0·980	0·993	1·006	1·017	1·027
(*e*)	0·978	0·978	0·982	0·987	0·993	0·998	1·004	1·009	1·014
Maximum	1·023	1·023	1·012	1·006	1·007	1·015	1·026	1·039	1·054

by H. A. David[3] and the other values have been obtained by fresh calculations. Considerable differences occur and no universal rule would appear to be applicable. In (*a*), (*b*) and (*c*) the value is always below one, whereas for (*d*) and (*e*) it starts below unity and changes over to above unity between n equal to seven and n equal to eight. It does, however, look as though in the majority of cases, and especially for small samples, the use of normal-curve factors will tend to underestimate the standard deviation. Increasing the size of sample is not necessarily a guarantee of increased accuracy in the estimation and in some cases can make the estimate even more inaccurate. Plackett[4] has shown that the ratio $w_n/d_n\sigma$ has a maximum expectation which is given in the last line of Table IV; in the sample sizes we are considering its highest value is plus 5%. For samples of size five and six the maximum is very little different from unity and it confirms that the most likely source of error for such sample sizes is the underestimation of the standard deviation, perhaps by as much as 10% in an extreme case. From Table III it is clear that an underestimate of standard deviation has a bigger effect on the probability level than an overestimate of the same amount would have.

Overall Effects on Chart for Means

We must now consider the effects that can arise if the samples are drawn from a non-normal population and normal-curve factors are used, first to estimate the standard deviation from the range and then to set up the limits for the means from the estimated standard deviation. The population used will be a χ^2 distribution with six degrees of freedom, for which $\gamma_1 = 1{\cdot}12$ and $\gamma_2 = 2$. Two different sample sizes, four and eight, will be considered. The standard deviation is estimated from the ranges in the samples by multiplying the mean range, $\overline{w}$, by $1/d_4$ or $1/d_8$ (0·485 7 or 0·351 2) respectively. This produces an underestimate of the population standard deviation, as the expected value of $\overline{w}/d_4$ is 0·961 times the standard deviation, and of $\overline{w}/d_8$ is 0·967 times

tandard deviation, of the χ^2 population from which the samples drawn. The sample mean will, however, be an unbiased estimate e population mean. Control limits can now be set up in the usual ner, but the expected proportion falling outside any control limit not be correct, because (*a*) the distribution of means is not normal (*b*) the standard deviation of the means has been underestimated. true probabilities of falling outside any stated limits can be com- d for these cases, and in Table V these true probabilities are given

TABLE V

True probabilities in calculated chart

inal el	Sample Size 4			Sample Size 8		
	Lower Tail	Upper Tail	Combined Tails	Lower Tail	Upper Tail	Combined Tails
	0·093 7	0·114 8	0·208 5	0·098 2	0·112 5	0·210 7
5	0·034 7	0·069 2	0·103 9	0·041 7	0·065 8	0·107 5
25	0·010 7	0·043 3	0·054 0	0·016 2	0·039 6	0·055 8
1	0·001 5	0·024 4	0·025 9	0·003 9	0·020 6	0·024 5
05	0·000 2	0·016 5	0·016 7	0·001 2	0·013 1	0·014 3

wide variety of nominal levels. Once again the individual tail abilities are often quite different from the nominal levels, but the ined tail probabilities are usually nearer. As the sample size is ased from four to eight the lower tail probabilities increase and pper ones decrease, but the combined tail position is not neces- an improvement. As the sample size increases still further, the tends to be more nearly normally distributed and the difference een the nominal and true probabilities is merely a reflection of rror in the estimation of standard deviation. The magnitudes of sulting changes in probability levels are then as shown in Table III. the systematic errors in estimation of standard deviation tend rally, but not invariably, to be smaller with larger samples it that we should use fewer large samples rather than more smaller les for the control chart on mean values if the nominal and true ol levels are to be kept as near as possible to one another.

on Chart for Range

e common quality control technique in setting up a chart for s is to find the mean value of the range from samples that are able and set up limits which are multiples of this mean range. The ples used are determined by reference to the distribution of ranges mples from a normal population. If in fact the population is not al and the usual multiples of mean range are used some sort of will be introduced. Cox[5] has shown, for sample sizes from two

to five and samples drawn from a variety of populations, how th coefficient of variation (i.e. the standard deviation/mean value) range varies from population to population. Over the range of valu of γ_2 most likely to be met with it appears that the coefficient of varia tion increases with γ_2, whereas the value of γ_1 does not appear to b related to the coefficient of variation. The effect of this will be tha with γ_2 less than zero the upper limits for range will be set too high b using normal curve factors and it will be difficult to detect small in creases in variability. Conversely, if the value of γ_2 is greater than zer the upper limits for range will be set too low, and more ranges will fa outside the limits than expected even with the process in control. Sinc it seems that the most likely forms of non-normal populations to b met with will have values of γ_2 greater than zero, this question o excessive numbers outside the limits must be guarded against. If ther is no knowledge of the parent population, excessive variability shoul first of all be investigated by making an examination of the shape o the parent population. On the other hand, if it is known from previou information that γ_2 is greater than zero it would be advisable to us control limits for range corresponding to fairly low percentage point It should be noted that increasing the sample size does not impro the situation, and, if anything, smaller samples taken more often ar better as they lead to less difference between the true and the norma percentage points of range, thus maintaining a more accurate chec on the variability of the process under consideration.

Conclusions

The general picture that emerges from this brief study of non normality is that the nominal significance levels in a control chart fo means should not be followed too slavishly unless there is evidence tha the basic distribution of measurements is effectively normal. Ther are grave dangers in utilising charts containing only an upper or lower significance level by itself, as even if the combined upper an lower levels give a probability level that is nearly correct the individua levels may be a very long way out. An increase in sample size wi effect some improvement but will always leave some error if the varia bility is estimated from the mean range in the available sample using the d_n factors from the normal distribution. There is therefor a strong case at the start of a scheme for estimating the variabili both by using the standard deviation in the samples and the range i the samples. If the two methods give similar results it is an indicatio that normal distribution theory is reasonably appropriate, but if the differ the standard deviation itself should be used and as large a sampl as is feasible should be employed for the chart.

Similarly in setting up the chart for ranges a comparison of the rang estimate of standard deviation with the actual standard deviatio should be made. If these agree it is reasonable to proceed in the usua manner, but if they disagree it is important to interpret the contro

for ranges in a broad sense only and not expect exactly the nal numbers of points to fall outside the various limits. Small les taken more often will produce the least discrepancy in this ct but as this would conflict with the requirement of the chart ieans a compromise must be sought in any practical situation.

should be noted that if a given number of observations N are to sed to set up a quality control chart, the random errors in the ation of the necessary quantities may depend on the particular od of subdivision of the N observations used. The estimation of rocess mean will have the same random errors whether we divide N observations into $\frac{1}{2}N$ samples of 2 or $\frac{1}{5}N$ samples of 5 or in any manner. However, the foregoing discussion has shown that for alculation of the control limits for the mean systematic errors are er for large sample sizes. When the standard deviation is to be ated from the range in a number of samples F. E. Grubbs and Weaver[6] showed that the optimum size of the samples into which N observations should be split in order to minimise the random s is about 7 or 8. The optimum size of the samples to keep the matic errors as small as possible appears to depend on the par- ar form of non-normal population from which the samples are n. From Table IV it appears, however, that samples of about 7 are likely to be the most reasonable size to use unless much is n about the form of the population being sampled. Thus a com- ise must be sought between the requirements from the point of of the means and from the point of view of the ranges. To divide N observations into a large number of very small samples can lead arked discrepancies in the chart for means unless the form of the lation is known and an appropriate allowance made in the cal- ion of the control limits.

REFERENCES

RSON, E. S. (1931). 'The analysis of variance in cases of non-normal variation', *Biometrika*, **23**, 114.

DENT' (1927). 'Errors of routine analysis', *Biometrika*, **19**, 151.

ID, H. A. (1954). 'The distribution of range in certain non-normal populations', *Biometrika*, **41**, 463.

KETT, R. L. (1947). 'Limits of the ratio of mean range to standard deviation', *Biometrika*, **34**, 120.

D. R. (1954). 'The mean and coefficient of variation of range in small samples from non-normal populations', *Biometrika*, **41**, 469.

BBS, F. E. and WEAVER, C. L. (1947). 'The best unbiased estimate of population standard deviation based on group ranges', *J. Amer. Statist. Ass.*, **42**, 224.

DIFFERENCES BETWEEN EXPERIENCED INTERVIEWERS

ALEXANDER BARR

Oxford Regional Hospital Board

In this article Mr Barr discusses the differences that aro between the results from two experienced interviewers in recent social enquiry. Although they agreed on matters fact, marked differences arose over the answers to questio involving an expression of opinion.

Investigating the differences in performance of experienced an inexperienced interviewers, Durbin and Stuart[1] concluded that th untrained interviewer was decidedly, though not overwhelmingl inferior to the professional person. Within each group of interviewe there was no marked heterogeneity between the individuals. The mai differences occurred between classes of interviewers rather than betwee individuals, and these differences were very little affected by the ci cumstances of the interview. In a continuation of the same analysi Booker and David[2] found no evidence for assuming that the resul recorded by the investigators arose from differing abilities or that th inexperience of the untrained persons led to their recording opinion preferences or facts significantly different from those recorded by ex perienced interviewers. In a more recent paper, Gray[3] indicated tha where the factual items were adequately defined, no significant inte viewer variability occurred. Contrariwise, in questions where a opinion was expressed there was appreciable variability which wa related to the interviewers' own opinions.

This paper briefly describes certain differences found between tw experienced interviewers in a recent social enquiry.

Material and Method

The aim of the survey was to discover what mothers in a certain tow thought of the existing maternity services and what additional facilitie were considered necessary. A random sample of mothers of babies bor in 1955 was selected from the records of the County Medical Officer o Health and two experienced interviewers were employed to visit th mothers in their homes. The interviewers knew each other personall having been employed for some years by the same Social Survey firm They had previously worked on the same investigating team and wer conversant with the procedure and the difficulties that can arise i surveys of public opinion. Each interviewer had the use of a car for th survey period. This minimised fatigue and at the same time expedite the collection of the information. The latter was important for admini trative reasons.

ıe addresses of all sample members were arranged in street order. e, in turn, were grouped into three fairly distinct geographical icts. In the first and smallest area the interviewers worked together oint and separate visits) with the object of establishing a common satisfactory routine. When this had been done, the interviewers menced working in the other two areas separately. These areas allocated randomly, and neither individual had any previous vledge of the district or town.

a special 'briefing session' held two days before the survey began, background and purpose of the survey were carefully outlined and rrangements for the field work were discussed. No written instruc- were provided, since only two interviewers were involved and it possible to keep in close touch with them as the survey progressed. interviewer was supplied with signed letters of introduction (on ed notepaper) and a copy was given or shown to each householder e necessary. No one refused to co-operate.

Questionnaire

ıe questionnaire was arranged in three parts; the first section con- d twenty-five questions of a general nature which were put to all ıers, and the second and third sections included special questions others confined in hospital and mothers confined at home respec- y. The principal questions asked are listed in the appendix.

ıe majority of questions were factual and usually could be answered ther of two alternatives. Thus, 'Was this your first baby?', 'Was irth normal?', 'Did you attend an ante-natal clinic?', 'Is this your house?', 'Was your husband on night duty when the baby was ?', and so on, could be answered by a direct 'Yes' or 'No'. Other ions, like 'Where was your baby born?', 'Where would you like to had your baby?', 'Where would you like to have your next baby?', l be answered by 'Home' or 'Hospital'. The questionnaire also ded other factual enquiries such as the age of other children in amily, number of rooms in the house, the length of time resident e area, and husband's occupation. In addition, it was necessary to de a number of questions of opinion; for example, 'How did you bout having booked at a hospital seven miles away?', and 'Why you have your baby at home/hospital?'. An analysis of replies ned by the two interviewers to some of the questions is given v.

ch individual question in the enquiry form was considered with the viewers at the 'briefing session', and where there was likely to be a rence of opinion the possible answers were discussed at length. As ılt, a number of differences between the interviewers were resolved everal minor amendments were made to the questionnaire. It was ıged that an assessment should be made at the end of the first day survey. This was duly carried out and no differences between the viewers were detected. As the interviewers were staying at the

same hotel it was easy for them to compare their results and discuss an particular problem that might arise during the course of the survey.

Results

Altogether, 258 mothers were interviewed over a period of a fortnigh and the number seen by each interviewer is shown in Table I. Th average number of successful interviews per working day for both inte viewers was 26. The proportion of home- and hospital-confine mothers falling to each was practically the same; this was fortuitou since no comprehensive information was available on this aspect befor the survey began. Of every five mothers contacted, roughly three ha been confined at home and two in hospital.

TABLE I

Distribution of sample members by place of confinement and by interviewer

Interviewer	Number Confined			Percentage Confined		
	Home	Hospital	All	Home	Hospital	All
A	84	50	134	62·7	37·3	100
B	79	45	124	63·7	36·3	100
Both	163	95	258	63·2	36·8	100

$\chi^2 = 0{\cdot}002,\ n = 1,\ 0{\cdot}95 > P > 0{\cdot}90$

The town is of comparatively recent development (a so-called 'Ne Town'), and the social and housing conditions, as shown by the 195 census, were above the general average for the country. The populatio is predominantly in the younger ages. Nearly all the employment i provided at the local steelworks and, as the executive staff lives outsid the town boundary, the town itself is virtually a 'one class' area. Ther are therefore good grounds for assuming that the sample group wa homogeneous in character. As a check, an analysis of the interviewe sample sub-groups by social class, age, parity, and housing condition was made, but no differences were revealed that could not easily hav arisen by chance.

The same number of multiparae (that is, women who have born previous children) was seen by each interviewer (Table II), the dif ference in the totals interviewed being accounted for by differen numbers of primiparae (that is, women bearing a first child). Of th 39 primiparae interviewed by A, 26 (66·7%) were confined in hospital while of the 95 multiparae, 24 (25·3%) were so confined. In the cas of B, 15 (51·7%) of the 29 primiparae and 30 (31·6%) of the 95 multi parae had their babies in hospital. Thus, on the whole, hospital con finements were commoner among primiparae than multiparae, an this is consistent with the trend throughout the country. Further whereas the proportion of hospital confinements among primiparae wa higher for A than for B, it was the opposite for multiparae. Th

rences, however, were not significant at the conventional 5% level robability. [$\chi^2 = 1{\cdot}64$, $n = 1$, $P \simeq 0{\cdot}2$. The value of χ^2 was derived llows. The separate values of χ^2 (each from a 2×2 table) expres- the difference between the proportion of hospital cases interviewed A and B among primiparae and multiparae were added, giving a ed χ^2 with 2 degrees of freedom. From this, the value of χ^2 for the rence between the proportions of hospital patients in A and B for iparae plus multiparae was subtracted, thus giving a χ^2 with 1 ee of freedom, which indicated the way the difference between A B changed between the primiparae and multiparae.]

TABLE II

Distribution of sample members by parity of mother and by interviewer

erviewer	Number of Mothers			Percentage		
	Primiparae	Multiparae	All	Primiparae	Multiparae	All
A	39	95	134	29·1	70·9	100
B	29	95	124	23·4	76·6	100
Both	68	190	258	26·7	73·3	100

$\chi^2 = 0{\cdot}80$, $n = 1$, $0{\cdot}50 > P > 0{\cdot}30$

ne mean age of the mothers and the average number of previous ren in the family were also similar for each observer's sample. efore, although it cannot be established definitely that there were ifferences between the two groups, no apparent bias was present. examining the composition of the sample it will be appreciated the analysis has been concerned with questions of fact only—place nfinement, parity of the mother, size of the family, etc. In each nce the agreement between the interviewers was found to be fairly . As a final illustration of the answers obtained to such questions, ttendances at the special health centre are given in Table III. The ortion attending was slightly higher in the case of B than of A, but ifferences were not statistically significant.

TABLE III

Distribution of sample members by attendance at health clinic and by interviewer

rviewer	Number of Mothers			Percentage		
	Attending Clinic	Not Attending Clinic	All	Attending Clinic	Not Attending Clinic	All
A	78	56	134	58·2	41·8	100
B	77	47	124	62·1	37·9	100
Both	155	103	258	60·1	39·9	100

$\chi^2 = 0{\cdot}26$, $n = 1$, $0{\cdot}70 > P > 0{\cdot}50$

The most vital questions in the enquiry concerned the reasons wh the mother had her baby at home or in hospital, as the case may be and the reaction of the mother who was confined at the local hospital Both of these were questions of opinion, and the 'prompt' answers ar shown in the appendix.

Table IV indicates the reasons for home or hospital confinemen Little difference occurred between the individual observers relative t hospital confinements, but the same could not be said about the con finements at home. The proportion found by observer B not going t hospital because it was too far away or because they had no option wa much greater than that found by observer A, and this difference wa very significant ($\chi^2 = 14{\cdot}05$, $n = 1$, $P < 0{\cdot}001$). But when this grou was combined with those who could not obtain admission to hospital thereby giving the potential number of mothers confined at home wh would rather have been in hospital, the difference between the investi gators was no longer significant ($\chi^2 = 2{\cdot}33$, $n = 1$, $0{\cdot}20 > P > 0{\cdot}10$) Hence it would seem that the proportion of women confined at hom who would have preferred confinement in hospital did not diffe between the two observer groups, whereas the main reasons for no going to hospital varied materially. The inference is that those wh

TABLE IV

Reason why mothers had their babies at home or in hospital

(*a*) *Home confinements*

Reason given by Mother	Interviewer			Percentage		
	A	B	Both	A	B	Both
Could not get into hospital..	28	15	43	33·3	19·0	26·4
Hospital too far away and had no option	12	33	45	14·3	41·8	27·6
Wanted to be near friends..	20	5	25	23·8	6·3	15·3
Previous births at home ..	13	6	19	15·5	7·6	11·7
Disliked hospital	8	6	14	9·5	7·6	8·6
Did not want to leave family	2	5	7	2·4	6·3	4·3
Considered medically unnecessary	—	4	4	—	5·1	2·5
Other	1	5	6	1·2	6·3	3·7
TOTAL ..	84	79	163	100	100	100

(*b*) *Hospital confinements*

Reason given by Mother	Interviewer			Percentage		
	A	B	Both	A	B	Both
Medical	28	27	55	57·1	58·7	57·9
Unsatisfactory home conditions	11	5	16	22·4	10·9	16·8
No one to look after house..	4	6	10	8·2	13·0	10·5
Other	6	8	14	12·2	17·4	14·7
TOTAL ..	49	46	95	100	100	100

could not obtain admission would have been willing to go to the existing hospital seven miles distant had suitable accommodation been available. The other group, however, would not have gone even if beds had been available, because it was considered that the hospital was too far. It will be clear that, other things being equal, a much stronger case for the erection of a new local hospital unit could be made out on the evidence supplied by B than that given by A.

As already mentioned, the nearest hospital was situated in an adjacent town seven miles away. A frequent bus service connects the two towns and travelling was not onerous. The 95 mothers having their babies in hospital were asked what they felt about being confined at a maternity hospital seven miles from their home, and the responses are given in Table V.

TABLE V

Reaction of mothers confined in hospital to booking at hospital seven miles distant. (Excluding five cases where no information was available)

Interviewer	Mothers' Reaction			Percentage		
	Didn't Mind	Disliked Idea	Total	Didn't Mind	Disliked Idea	Total
A	30	18	48	62·5	37·5	100
B	12	30	42	28·6	71·4	100
Both	42	48	90	46·7	53·3	100

$\chi^2 = 9{\cdot}04,\ n = 1,\ 0{\cdot}01 > P > 0{\cdot}001$

In this respect one observer was practically the complete antithesis of the other, and the results were similar for primiparae and multiparae separately. It will be noted that observer B had the higher proportion of mothers who disliked the idea of going so far, coinciding with the higher proportion of the mothers confined at home for the same reason.

Discussion

The purpose of the survey was adequately explained to the interviewers at a meeting preceding the field work. In particular, the survey set out to ascertain if there were any special grounds for providing additional maternity facilities either in the form of an extension to the present hospital, which was seven miles distant from the town surveyed, or by the building of a new maternity unit in the town itself. Naturally, local publicity had been very vigorous, and it was expected that opinion might be coloured by this activity. Both interviewers employed in the survey were equally aware of this factor. Albeit there were marked differences between them in certain essential points. When the interviewers were asked for their reactions to the survey afterwards, observer B, whose sample mothers showed a leaning towards a general dissatisfaction with the present arrangements, stated that she personally had been quite indifferent to the survey at the beginning but, as the

survey progressed, she became thoroughly convinced that a new unit was required in the town. The other interviewer believed that there might be a case for some extra facilities, but was sceptical about the need for any new building. The conclusions to be drawn from this analysis are similar to and confirm those already reported by Gray.[3] As far as factual information was concerned there was little difference between the two observers, but substantial variations occurred in regard to opinion questions.

Although particular attention was paid to the compilation of the 'prompt' answers to opinion questions, wide variation in the interpretation took place. It is often impossible to anticipate every answer that is likely to be given, and a trial run is obviously desirable. It is quite conceivable that the order in which the prompts are arranged has some influence on the ultimate results, but this survey did not lend itself to investigating this matter. As it was, the prompt answers were tabulated in what was thought to be the order of importance, and they were put to the mothers in this sequence. This may have resulted in an overstatement of the initial answers, because the mother, having selected the first or second reason, may not have been aware or conscious of the other reasons, which were consequently ignored. It might have been advantageous to have arranged the prompt answers in reverse order of importance with the object of sustaining the attention. While this may possibly have accounted for the differences in some of the less common reasons (for example 'disliked hospital' and 'did not want to leave family' in the case of home confinements), it does not account for the difference between the observers in the proportions stating that the hospital was too far away.

In trying to account for the interviewer differences it must be remembered that the subject of the survey possessed some emotional feeling, but this is not unusual in investigations of a social nature. There is always a danger that even experienced interviewers will be influenced by local propaganda and outlook, and it seems worth while taking precautionary measures in this direction, especially if the sample and the number of interviewers involved are small. One possible method of keeping a concurrent check on the results would be to provide each interviewer with a specially designed 'Way Bill' on which each day's interviews would be recorded analytically. This should provide the basis for a simple and expeditious analysis of all the important aspects of the survey at frequent intervals. Immediate measures could then be taken to correct any untoward differences that might appear between interviewers instead of leaving them to be discovered after completion of the survey. An extension which might be worth incorporating would be for interviewers to make a number of joint visits periodically throughout the survey, thereby providing opportunity for checking each other's standard and approach. Obviously the latter proposal is only practical where the number of interviewers is small and the survey area compact.

APPENDIX: QUESTIONNAIRE

Summary of Principal Questions

A. *All Mothers*

Question	*Prompt Answers*	
Where did you have your baby?	Home	Hospital
If hospital, please give name of hospital ..		
If not at home or hospital, please give address		
Was birth normal?	Yes	No
Was this your first baby?	Yes	No
If not first baby, how many other children in family?		
Please give ages and where born	Home	Hospital
Did you go to an Ante-Natal Clinic before present birth?	Yes	No
How long have you lived in the town? ..		Years
Where did you live previously?		
Is this your own house?	Yes	No
How many rooms have you?	Living rooms/Bedrooms	
Does anyone else live with you (e.g. friends, relations, boarder)?	Yes	No
Where does your husband work?		
What is his occupation?		
Did you work before your baby was born? ..	Yes	No
If so, when did you stop working?		

B. *Mothers confined in Hospital*

Why did you have your baby in hospital? ..	Medical Unsatisfactory home conditions No one to look after house Financial reasons Previous children born in hospital Other (please state)
Would you have preferred to have had your baby at home if it could have been arranged to your satisfaction?	Yes No
How were you transported to hospital? ..	By private car By taxi By ambulance
Did you have any difficulty over transport? ..	Yes No
How did you feel about having booked at a hospital seven miles away?	Didn't mind Disliked the idea Other (please state)
Did you have anyone to look after the house while you were in hospital?	Yes No

B. Mothers confined in Hospital (continued)

Question	*Prompt Answers*	
If so, was this person your	Mother (in-law) Sister Other family relative Friend	
Were your husband and friends able to visit you in hospital regularly and easily? ..	Yes	No
If you have another child would you like to have it at home or in hospital?	Home	Hospital

C. Mothers confined at Home

Why did you have your baby at home? ..	Didn't like hospital Wanted to be near friends Could not make arrangements with hospital Hospital too far away Previous children born at home Other (please state)	
Would you have preferred to have had your baby in hospital?	Yes	No
Would you have preferred to have had your baby in hospital if you could have chosen your length of stay?	Yes	No
Did you book a doctor?	Yes	No
Did you book a midwife?	Yes	No
Did you have any member of your family to help you?	Yes	No
If so, was this person your	Mother (in-law) Sister Other family relative Friend	
If you have another child would you like to have it at home or in hospital?	Home	Hospital

REFERENCES

[1] Durbin, J. and Stuart, A. (1951). 'Differences in response rates of experienced and inexperienced interviewers', *J. R. Statist. Soc.*, A, **114**, 163.

[2] Booker, H. S. and David, S. T. (1952). 'Differences in results obtained by experienced and inexperienced interviewers', *J. R. Statist. Soc.*, A, **115**, 232

[3] Gray, P. G. (1956). 'Examples of interviewer variability taken from two sample surveys', *Applied Statistics*, **5**, 73.

'THREE-GROUP' REGRESSION ANALYSIS*

PART II. MULTIPLE REGRESSION ANALYSIS

WENDY M. GIBSON and GEOFFREY H. JOWETT
University of Sheffield

In an earlier article Miss Gibson and Dr Jowett described some applications of the 'three-group' method to simple regression analysis. The present article shows how the method can be extended to give a multiple regression equation with two independent variables, together with estimates of the residual scatter and the standard errors of the regression coefficients.

The three-group method of regression analysis, which we described in the last issue of this journal,[1] is readily generalised into a technique for multiple regression. In the simplest case, when there are only two independent variables, the method is arithmetically easy, and also elegant and illuminating. It becomes increasingly complicated as further independent variables are introduced; but workers who are expert enough to contemplate an analysis involving three or more independent variables will probably prefer in any event to use the method of least squares. This paper has the dual aim of showing how the three-group method may be applied in the simplest case and of introducing readers unfamiliar with the technique to some of the ideas of multiple regression.

The usual expositions of multiple regression analysis with two independent variables, of which an excellent example was given by Professor E. S. Pearson,[3] involve the notion of a *regression plane* which is fitted by least squares to the observations. An alternative exposition in non-geometrical terms which is very favourably received by, say, accountants, and which is especially suited to the use of three-group methods, is as follows.

Suppose we have a dependent (or 'effect') variable z affected by two independent (or 'cause') variables, x and y, which tend to vary together (possibly for some causally irrelevant reason) in a particular set of data. For example, in the blast furnace data of Table I, the output of pig iron per ton of coke, z, which is a measure of operating efficiency, is affected for quite different reasons by the percentage of available carbon in the coke x (because it is largely the carbon which does the work of smelting), and by the hearth temperature of the furnace. The latter cannot be measured directly but is believed to be reflected, other

* Read in part to the Sheffield Group of the Industrial Applications Section of the Royal Statistical Society, October 1956.

TABLE I

Blast furnace data; coded weekly records of output of pig iron per ton of coke (z), available carbon in coke (x), and silicon in pig iron (y)

Serial No.	x	y	z	Group	Serial No.	x	y	z	Group
1	50	12	56	UL	38	24	47	48	LU
2	52	15	59	UL	39	22	44	43	LM
3	31	26	66	ML	40	18	66	10	LU
4	43	50	65	UU	41	19	70	34	LU
5	41	35	61	MM	42	13	54	35	LU
6	41	24	59	ML	43	23	24	45	LL
7	43	36	59	UM	44	14	48	40	LU
8	38	53	59	MU	45	19	74	14	LU
9	51	33	66	UM	46	18	41	6	LM
10	54	34	97	UM	47	15	67	14	LU
11	39	32	90	ML	48	12	63	16	LU
12	45	34	35	UM	49	28	45	32	MM
13	35	41	62	MM	50	47	47	21	UU
14	44	18	63	UL	51	11	34	61	LM
15	56	19	67	UL	52	14	35	70	LM
16	54	34	49	UM	53	33	23	45	ML
17	49	55	41	UU	54	26	22	45	LL
18	51	42	39	UM	55	48	28	71	UL
19	38	40	55	MM	56	33	12	67	ML
20	33	41	74	MM	57	28	27	63	ML
21	28	53	75	MU	58	19	33	59	LM
22	36	31	59	ML	59	33	35	61	MM
23	38	41	76	MM	60	28	43	51	MM
24	8	45	64	LM	61	46	37	44	UM
25	8	43	16	LM	62	48	25	56	UL
26	8	42	62	LM	63	36	20	57	ML
27	26	40	50	LM	64	16	48	36	LU
28	29	35	70	MM	65	34	50	40	MU
29	44	26	65	UL	66	35	37	44	MM
30	38	35	76	MM	67	37	45	63	MM
31	43	41	53	UM	68	53	62	63	UU
32	39	47	46	MU	69	41	70	73	MU
33	49	42	59	UM	70	34	30	75	ML
34	52	28	53	UL	71	27	47	46	LU
35	35	47	48	MU	72	31	21	67	ML
36	28	50	43	MU	73	36	26	70	ML
37	14	35	57	LM	74	63	35	54	UM

things being equal, in the lowness of the ultimate percentage, y, of silicon in the pig iron coming out of the furnace. However, at the particular coking plant where the coke in this example was made there was a tendency throughout the seventy-four weeks of the investigation for high values of x to be associated with strength and porosity in the coke—qualities which both tend to induce a high hearth temperature and therefore a low value of y. The negative correlation between x and y was due to conditions at the coking plant, and could presumably be changed (or so we will assume) without altering the direct effects of x and y on z; the direct 'effect' of y is really, of course, a measurable reflection of the direct effect of hearth temperature.

Hence the apparent effect, on the average, of a unit increase in x on z, which is measured by the regression coefficient b_{zx}, can be split into two parts:

(i) *direct effect* (in this case due to chemical energy)

ii) *indirect effect* due to a tendency (the causal mechanism of which is not in question) for lower values of y to occur with higher values of x, the observed average difference in y per unit of x being measured by b_{yx}.

If, using Yule's notation, we describe the direct effects of unit increases in x and y on z by the partial regression coefficients $b_{zx.y}$ (i.e. the hypothetical effect of a unit increase in x on z, y being held constant meanwhile) and $b_{zy.x}$, we may use (i) and (ii) to account for the whole of b_{zx} as in equation (1):

$$\underset{\substack{\text{apparent effect}\\ \text{of } x \text{ on } z}}{b_{zx}} = \underset{\substack{\text{direct effect}\\ \text{of } x \text{ on } z}}{b_{zx.y}} + \underset{\substack{\text{direct effect of a}\\ \text{unit change in } y \text{ on } z}}{b_{yx}\overset{\substack{\text{average change of}\\ y \text{ with } x \text{ in data}}}{b_{zy.x}}} \qquad \ldots.(1)$$

By a similar argument, starting from y, we may deduce a second equation:

$$b_{zy} = b_{zy.x} + b_{xy}b_{zx.y} \qquad \ldots.(2)$$

To estimate the direct effects of x and y on z the ordinary regression coefficients in (1) and (2) are calculated direct from the data by three-group methods or otherwise, and the equations are then solved simultaneously for $b_{zx.y}$, $b_{zy.x}$. We note in passing that (1) and (2) reduce to the ordinary normal equations if the ordinary regression coefficients are estimated by least squares.

Solving (1) and (2):

$$b_{zx.y} = (b_{zx} - b_{yx}b_{zy})/(1 - b_{xy}b_{yx}) \qquad \ldots.(3)$$

$$b_{zy.x} = (b_{zy} - b_{xy}b_{zx})/(1 - b_{xy}b_{yx}) \qquad \ldots.(4)$$

The amount by which z may be expected to be above average for any combination of x and y is then given by (5), which is the equation of the plane of regression:

$$\mathrm{E}(z - \bar{z}) = b_{zx.y}(x - \bar{x}) + b_{zy.x}(y - \bar{y}) \qquad \ldots.(5)$$

where $\bar{x}$, $\bar{y}$, $\bar{z}$ are the average values of x, y, z in the data, and E is an operator denoting expected value.

Example of Determination of $b_{zx.y}$ and $b_{zy.x}$ by Three-group Method

First a scatter diagram for x and y is drawn, as shown in Fig. 1, and beside each point is written a serial number for identification in Table 1, along with the corresponding z-value (preferably in different colours to avoid confusion). This diagram is very important in itself, showing

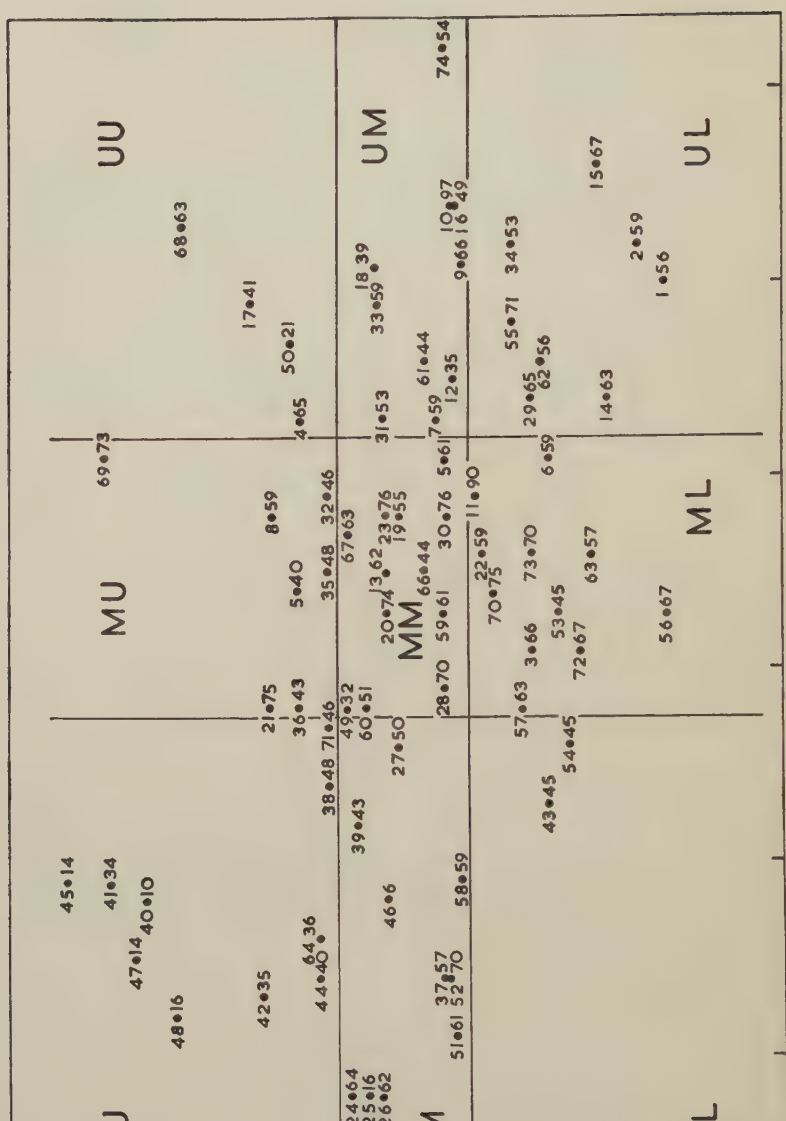


Fig. 1. Blast furnace data: scatter diagram divided into sections. The z-values are shown on the right of the points and the identi-

the extent to which high values of x and y are associated in the data and what combinations of values tend to lead to high values of z. Lines are then drawn parallel to the x- and y-axes dividing the scatter diagram into sections, as for simple regression on x or y. This creates nine groups, which are conveniently denoted by LU, LM, LL, etc.

s shown in Fig. 1, where LU, for example, signifies the group with ›wer values of x and upper values of y. Some of the nine groups may e empty, but this does not matter.

The symbol of the group to which it belongs is written beside each ntry on the data sheet, as in Table I. The totals of x, y, and z in each f the nine groups are then calculated, as in Table II, and checked

TABLE II

Blast furnace data: computation of partial regression coefficients

Group	Totals			n	Means		
	x	y	z		x	y	z
LL	49	46	90	2	—	—	—
LM	148	392	488	10	—	—	—
LU	177	584	293	10	—	—	—
ML	378	272	718	11	—	—	—
MM	413	473	725	12	—	—	—
MU	243	370	384	7	—	—	—
UL	392	171	490	8	—	—	—
UM	499	368	555	10	—	—	—
UU	192	218	190	4	—	—	—
Total	2 491	2 894	3 933	74	33·663	39·108	53·149
L·	374	1 022	871	22	17·00	46·45	39·59
U·	1 083	757	1 235	22	49·23	34·41	56·14
·L	819	489	1 298	21	39·00	23·29	61·81
·U	612	1 172	867	21	29·14	55·81	41·29
Sum	2 888	3 440	4 271	86	—	—	—
U· − L·	—	—	—	—	32·23	−12·04	16·55
·U − ·L	—	—	—	—	−9·86	32·52	−20·52

$b_{yx} = -0{\cdot}374$; $b_{xy} = -0{\cdot}303$; $b_{zx} = 0{\cdot}513$; $b_{zy} = -0{\cdot}631$
$(1 - b_{yx}b_{xy}) = 0{\cdot}887$; $b_{zx.y} = 0{\cdot}312$; $b_{zy.x} = -0{\cdot}536$

gainst the grand totals. Next the group totals are combined as follows using a capital letter to represent a total):

$$X_{\mathrm{L}\cdot} = X_{\mathrm{LU}} + X_{\mathrm{LM}} + X_{\mathrm{LL}} \qquad \ldots.(6)$$

$$X_{\cdot\mathrm{L}} = X_{\mathrm{UL}} + X_{\mathrm{ML}} + X_{\mathrm{LL}} \quad \text{etc.} \qquad \ldots.(7)$$

useful check on this totalling is given by such identities as

$$X_{\mathrm{L}\cdot} + X_{\mathrm{U}\cdot} + X_{\cdot\mathrm{L}} + X_{\cdot\mathrm{U}} = X_{\cdot\cdot} + X_{\mathrm{LL}} + X_{\mathrm{LU}} + X_{\mathrm{UL}} + X_{\mathrm{UU}} - X_{\mathrm{MM}} \qquad \ldots.(8)$$

here $X_{\cdot\cdot}$ is the sum of all the x values.

The corresponding means are then calculated, and from these the otal regression coefficients are deduced as described in the previous aper; typically,

$$b_{zx} = (\bar{z}_{\mathrm{U}\cdot} - \bar{z}_{\mathrm{L}\cdot})/(\bar{x}_{\mathrm{U}\cdot} - \bar{x}_{\mathrm{L}\cdot}) \qquad \ldots.(9)$$

inally, $b_{zx.y}$, $b_{zy.x}$ are calculated, using (3) and (4); for the blast rnace data their values are shown at the foot of Table II.

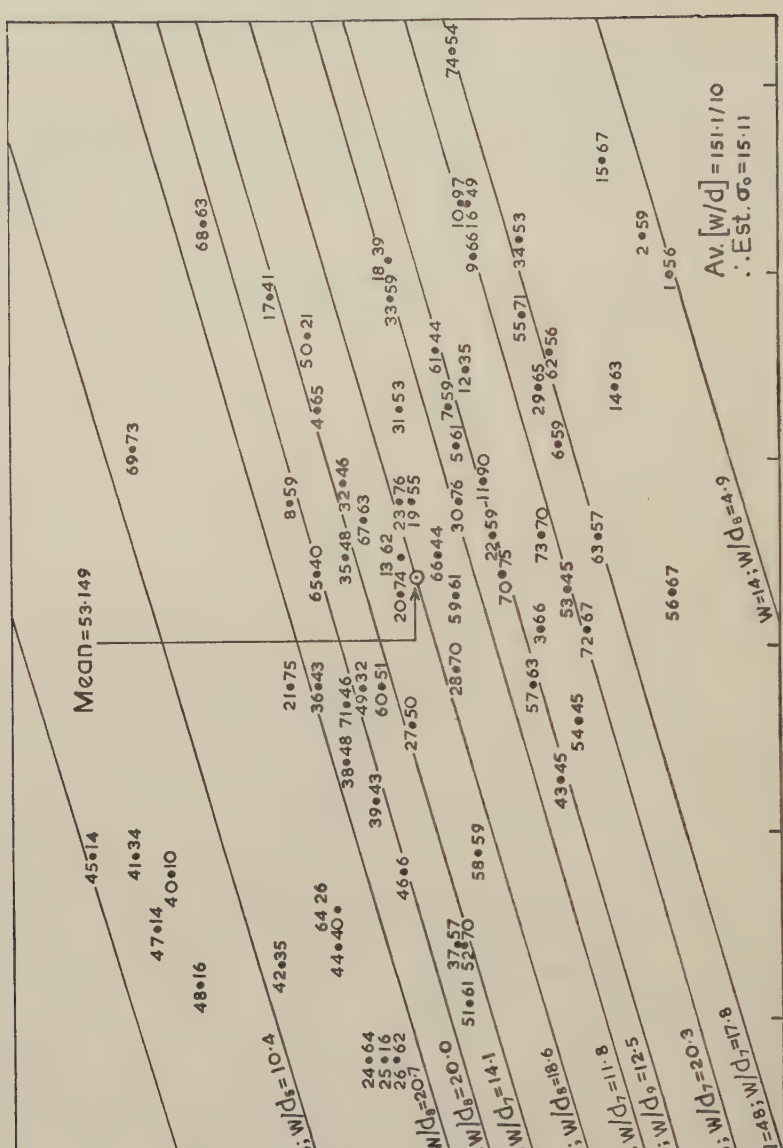


Fig. 2. Blast furnace data: scatter diagram with contour bands for estimating σ_0. The z-values are shown on the right of the

Measuring Residual Scatter

The first step in the determination of the standard errors of $b_{zx.y}$ $b_{zy.x}$ is to estimate the standard deviation σ_0 of the residual variation (i.e. variation unexplained in terms of x and y) of z. This is most conveniently done by dividing the observations into fairly equal groups within which the joint effect of x and y can be taken as roughly constant, and deducing an estimate of σ_0 from the ranges of z within these

roups. This is analogous to the method proposed for simple regression; it is accomplished in practice by dividing up the scatter diagram nto bands (preferably containing roughly equal numbers of points, nd fewer than twelve) by means of parallel lines with $(-b_{zx.y}/b_{zy.x})$ s slope. In practice this would be done on Fig. 1, but we have shown t in a separate diagram, Fig. 2, to avoid confusion. Although the fact s not immediately relevant, it is useful to note that these lines are contours of the plane of regression; the line through $(\bar{x}, \bar{y})$ is the contour corresponding to $\mathrm{E}(z) = \bar{z}$, where E denotes expected value, nd that meeting the y-axis k units further up is the contour corresponding to

$$\mathrm{E}(z) = \bar{z} + kb_{zy.x} \qquad \ldots.(10)$$

The ranges, w, of z-values for the points in each band are read off the diagram, and after being divided by the appropriate conversion factor l are averaged to give an estimate of σ_0. Some of the bands will be wider than others, and may tend to give greater ranges for that reason; f this is considered to be serious enough to matter, the remedy lies in ubdividing each band that is too wide into, say, two narrower bands nd using the average of the two sub-ranges instead. This estimation nas been carried out for the blast furnace data on the left of Fig. 2, the final estimate of σ_0 being shown in the bottom right-hand corner. At his stage of analysis it may be profitable to examine Fig. 2 critically or trends in z-values along the bands (e.g. tendencies to drop in the middle) which are indications of non-linearity in the effects of x and y, nd of the need for further analysis.

Estimation of the Standard Errors of $b_{zy.x}$ and $b_{zx.y}$

Since, for the purposes of calculating standard errors, b_{yx} and b_{xy} may be treated as constants, it follows directly from (3) that

$$\mathrm{var}(b_{zx.y}) = [\mathrm{var}(b_{zx}) - 2b_{yx}\mathrm{cov}(b_{zx}, b_{zy}) + b^2_{yx}\mathrm{var}(b_{zy})]/(1 - b_{xy}b_{yx})^2 \qquad \ldots.(11)$$

nd similarly from (4) that

$$\mathrm{var}(b_{zy.x}) = [\mathrm{var}(b_{zy}) - 2b_{xy}\mathrm{cov}(b_{zx}, b_{zy}) + b^2_{xy}\mathrm{var}(b_{zx})]/(1 - b_{xy}b_{yx})^2 \qquad \ldots.(12)$$

To estimate $\mathrm{var}(b_{zy})$, $\mathrm{cov}(b_{zx}, b_{zy})$, and $\mathrm{var}(b_{zx})$ we generalise the method described in the previous paper. Taking

$$z = \alpha + \beta_1 x + \beta_2 y + \xi$$

where β_1, β_2 are the 'true' values of $b_{zx.y}$ and $b_{zy.x}$, and ξ is the residual error of z, having standard deviation σ_0:

$$\begin{aligned} b_{zy} &= (\bar{z}_{.\mathrm{U}} - \bar{z}_{.\mathrm{L}})/(\bar{y}_{.\mathrm{U}} - \bar{y}_{.\mathrm{L}}) \\ &= [(\alpha + \beta_1\bar{x}_{.\mathrm{U}} + \beta_2\bar{y}_{.\mathrm{U}} + \bar{\xi}_{.\mathrm{U}}) \\ &\qquad - (\alpha + \beta_1\bar{x}_{.\mathrm{L}} + \beta_2\bar{y}_{.\mathrm{L}} + \bar{\xi}_{.\mathrm{L}})]/(\bar{y}_{.\mathrm{U}} - \bar{y}_{.\mathrm{L}}) \\ &= \beta_1 b_{xy} + \beta_2 + (\bar{\xi}_{.\mathrm{U}} - \bar{\xi}_{.\mathrm{L}})/(\bar{y}_{.\mathrm{U}} - \bar{y}_{.\mathrm{L}}) \end{aligned}$$

Since β_1, β_2, b_{xy} are constants:

$$\begin{aligned}\text{var}(b_{zy}) &= \text{var}(\bar{\xi}_{.U} - \bar{\xi}_{.L})/(\bar{y}_{.U} - \bar{y}_{.L})^2 \\ &= (1/n_{.U} + 1/n_{.L})\sigma_0^2/(\bar{y}_{.U} - \bar{y}_{.L})^2 \qquad \ldots\ldots(13)\end{aligned}$$

Similarly:

$$\text{var}(b_{zx}) = (1/n_{U.} + 1/n_{L.})\sigma_0^2/(\bar{x}_{U.} - \bar{x}_{L.})^2 \qquad \ldots\ldots(14)$$

Also:

$$\begin{aligned}\text{cov}(b_{zx}, b_{zy}) &= \text{cov}[(\bar{z}_{U.} - \bar{z}_{L.}), (\bar{z}_{.U} - \bar{z}_{.L})]/(\bar{x}_{U.} - \bar{x}_{L.})(\bar{y}_{.U} - \bar{y}_{.L}) \\ &= [\text{cov}(\bar{z}_{U.}, \bar{z}_{.U}) + \text{cov}(\bar{z}_{L.}, \bar{z}_{.L}) - \text{cov}(\bar{z}_{L.}, \bar{z}_{.U}) \\ &\quad - \text{cov}(\bar{z}_{U.}, \bar{z}_{.L})]/(\bar{x}_{U.} - \bar{x}_{L.})(\bar{y}_{.U} - \bar{y}_{.L}) \\ &= [(n_{UU}/n_{U.}n_{.U}) + (n_{LL}/n_{L.}n_{.L}) + (n_{LU}/n_{L.}n_{.U}) \\ &\quad + (n_{UL}/n_{U.}n_{.L})]\sigma_0^2/(\bar{x}_{U.} - \bar{x}_{L.})(\bar{y}_{.U} - \bar{y}_{.L}) \qquad \ldots\ldots(15)\end{aligned}$$

These formulae, though complicated in appearance, are easily worked out, and the resulting quantities substituted in (11) and (12). The computational layout for the blast furnace data is shown in Table III.

TABLE III

Blast furnace data: sampling variances and covariances of partial regression coefficients

	For-mula	Multiplier of σ_0^2 in Formula			√(Num./Den.)	Stan-dard Errors
		Numerator	Denominator	Num./Den.		
$\text{var}(b_{zx})$	(14)	0·090 9	1 038·77	0·000 087 5	—	—
$\text{var}(b_{zy})$	(13)	0·095 2	1 057·55	0·000 090 0	—	—
$\text{cov}(b_{zx}, b_{zy})$	(15)	−0.026 0	1 048·12	−0·000 024 8	—	—
$\text{var}(b_{zx.y})$	(11)	0·000 081 5	0·786 8	0·000 103 6	0·010 2	0·154
$\text{var}(b_{zy.x})$	(12)	0·000 083 0	0·786 8	0·000 105 5	0·010 3	0·156

TABLE IV

Three-group and least-squares results

	Three-group Estimates	Least-squares Estimates
b_{yx}	−0·374	−0·346
b_{xy}	−0·303	−0·329
b_{zx}	+0·513	+0·511
b_{zy}	−0·631	−0·616
Est. σ_0	15·1	15·4
$b_{zx.y}$	+0·312 (S.E. 0·154)	+0·337 (S.E. 0·143)
$b_{zy.x}$	−0·536 (S.E. 0·156)	−0·505 (S.E. 0·138)

S.E. signifies standard error.

Results obtained in this way and, for comparison, by the ordinary least-squares method are summarised in Table IV, where it will be seen that the accuracy of the three-group coefficients compares very favourably with that of the least-squares coefficients. Actually the estimate of σ_0 turned out to be slightly smaller than that obtained by least squares, which tends to make the standard errors smaller in proportion, but allowing for this does not in fact make much difference. Both coefficients would be adjudged significant on the five per cent level, using twice the standard error as the criterion of significance.

When this regression was calculated it was believed that choice of the optimum ratios (see the previous paper[1]) mattered more than subsequently proved to be the case, and since the distributions of x and y had rather short tails, the ratios 3 : 4 : 3 were taken for the groups instead of the generally recommended ratios 1 : 2 : 1.

Since the data in this particular example are in the form of time series, with noticeable correlation between successive observations, these results must be treated with reserve, and should be regarded merely as an illustration of the application of the method. For their correct analysis some prior trend removal (cf. Jowett[2]) would seem to be strictly desirable. Also, our explanation of the causal system operating has been simplified for purposes of exposition.

Acknowledgement

We wish to thank the United Steel Companies Ltd., Workington, for permission to use the blast furnace data.

REFERENCES

[1] Gibson, W. M. and Jowett, G. H. (1957). 'Three-group regression analysis. Part I. Simple regression analysis', *Applied Statistics*, **6**, 114.

[2] Jowett, G. H. (1955). 'Least squares regression analysis for trend-reduced time series', *J. R. Statist. Soc.*, B, **17**, 91.

[3] Pearson, E. S. (1936). Contribution to the discussion following a paper by W. J. Jennett and B. P. Dudding. 'The application of statistical principles to an industrial problem', *Suppl. J. R. Statist. Soc.*, **3**, 1.

A SIMPLE SEQUENTIAL PROCEDURE TO TEST WHETHER AVERAGE CONDITIONS ACHIEVE A CERTAIN STANDARD

ROLFE C. TOMLINSON

Scientific Department, National Coal Board

In this article Mr Tomlinson describes a sequential test which is easy to apply, reasonably efficient, and can be used when the distribution of sample values is not normal.

The Practical Problem

A large number of industrial occupations have associated with them a disease which arises from the environment in which the men work. Most of these diseases are cumulative in effect, in that men who work in the environment for only a short time do not suffer serious ill effects, and that the onset of the disease appears to depend on how badly the environment is contaminated and the length of exposure of the man to it. Possibly the most serious industrial disease of this kind in the coalmining industry is pneumoconiosis. This disease arises from the presence of dust in the lung and, in its advanced stages, leads to a progressive deterioration of the lung. The exact mechanism by which pneumoconiosis starts and develops is as yet imperfectly understood; indeed, a research programme to last ten to fifteen years has recently been commenced by the National Coal Board in an attempt to obtain fuller information.

The simplest hypothesis with regard to the occurrence of pneumoconiosis is that a man's liability to the disease is a simple function of the total amount of dust consumed during his working life, i.e. of his total exposure to dust. If this hypothesis is true it means that in order to reduce the chance of a worker contracting the disease in the course of his working life it is necessary to ensure that his total exposure during that period is below a certain level. This can only be ensured by setting a limit to the average dust concentration in the environment in which the man works. Within reason, it will not matter if conditions at certain times are worse than the average, provided that such periods are counteracted by corresponding periods of low dust concentration. Although the desirable average level of concentration for different types of dust cannot be fixed precisely until more research has been carried out, certain standards of dustiness have been laid down on the basis of past experience, and wherever possible the working places of all men employed in British coal mines are made to comply with these

approved conditions. In order to do this the industry is faced both with engineering problems of dust suppression and also with the problem of devising and carrying out a sampling scheme which will tell the engineers whether approved conditions have or have not been attained. The present paper is concerned with the problem of devising such a sampling scheme.

The natural variation in the dust concentration of air on a coal face is very large. Even when a single sample represents average conditions over a period of about half an hour, which is normally the case, samples taken on the same working shift may vary by as much as ten to one in measured concentration. Average conditions from shift to shift may vary by as much as two to one even when occasional 'rogue' results are excluded. Because of this, it is necessary to consider the average dust concentration over a shift, obtained from a sample spread throughout the shift, and sampling must be carried out on a number of shifts before it is possible to decide whether conditions are approved or not. As there are seldom more than twenty men at a working place it is obvious that unless the number of shifts sampled is reduced to a minimum, the number of samplers employed will be an appreciable proportion of the number of men carrying out productive work. It is essential, therefore, to provide a solution to the following simple problem:

> 'On how many shifts must we sample before we can state whether the average conditions over a period are above or below the approved value?'

The conventional answer to the question might be to state that each face must be sampled a fixed number of times, but this would be an unwise and wasteful solution. In view of the need for economy it appears to be important to adopt a sequential procedure that will enable one to stop sampling as soon as sufficient information has been obtained; this should reduce the number of samples required to a minimum. It should be noted, however, that the conditions under which this sequential plan must be applied differ considerably from those under which sequential methods are usually operated. Whereas sequential procedures are most often applied to the sampling of manufactured products, the average sample number being generally of the order of twenty or thirty, in the present case the variable being measured is under much less control than a manufactured product, and the average sample number required is of the order of two or three.

The Statistical Problem

The problem given above can be simply stated in statistical terms as follows:

Given a population with *median* m, and standard deviation σ, we wish to discover, by examining the results after each sample is taken, whether m is greater or less than a critical value μ. As soon as such a

decision is taken sampling will cease. In many cases it may be assumed that σ is known, but we shall also consider the case where σ is unknown. [In this paper it is convenient to use the median (the point such that 50% of values lie on either side of it) rather than the mean. However, for normal distributions, and indeed for all symmetrical distributions, the two are identical.]

The choice of scheme depends on the chance that can be accepted of making a wrong decision, i.e. of accepting a face when $m > \mu$, or of rejecting a face when $m < \mu$. In the present case we shall assume that there should be a 50/50 chance of accepting (or rejecting) a face when $m = \mu$, and only a 1 in 1 000 chance (approximately) of making the wrong decision when $|m - \mu| = 2\sigma$, when the distribution is normal, or more generally when $m = \lambda_u$ and λ_l where λ_u and λ_l are the upper and lower 95% points of the distribution when $m = \mu$.

Translating this statistical problem into practical terms, a single sample represents the average dust concentration over a single shift, the population median m represents the average dust concentration over a period, and the critical value, μ, is the 'approved' standard.

In seeking a solution to this statistical problem it is important to realise that we are assuming an idealisation of actual conditions. It assumes, as all sampling schemes must, that the conditions when samples taken are typical of conditions over a period; if conditions are quite atypical, then no sampling scheme will provide a satisfactory answer, and if such exceptional cases occur frequently the only satisfactory solution is continuous recording of dust conditions. A more serious problem arises from the fact that working conditions underground never remain similar from day to day, and that over a period of time there is likely to be a tendency for concentrations to alter because of changes in mining conditions. Thus although the scheme, if applied rigidly, might imply that a working place could be sampled two or three times and never again, it is necessary in practice to resample at intervals to ensure that conditions have not changed. Nevertheless, such a scheme does ensure, so far as any scheme can, that the amount of sampling is reduced to the minimum necessary to answer the basic question whether the working place is approved or not.

A Conventional Test

The conventional method of solving this problem was originated by Wald and has been described in many text books (e.g. *The Design and Analysis of Industrial Experiments*[1]). The procedure is to plot the cumulative totals after successive observations on a chart, such as Fig. 1, which has two boundary lines on it. When the cumulative total crosses one of the boundary lines, conditions are said to be 'approved' or 'not approved' as the case may be. The position of the boundary lines depends on the conditions specified for the test. If we take as our basic conditions that the sample values are distributed normally, and that there should be a probability α (in the case already mentioned

in 1 000) of rejecting a face when $m = \mu - 2\sigma$ or of accepting a ace when $m = \mu + 2\sigma$, the limiting lines would then by given by:

$$T_l = n\mu - \frac{\sigma}{4}\log_e\left(\frac{1-\alpha}{\alpha}\right)$$

nd

$$T_u = n\mu + \frac{\sigma}{4}\log_e\left(\frac{1-\alpha}{\alpha}\right)$$

where n is the number of samples. Thus in a particular case we have

$$\alpha = 0{\cdot}001,\ \mu = 100, \text{ and } \sigma = 20$$
$$T_l = 100n - 34$$
$$T_u = 100n + 34$$

These lines have been drawn in Fig. 1.

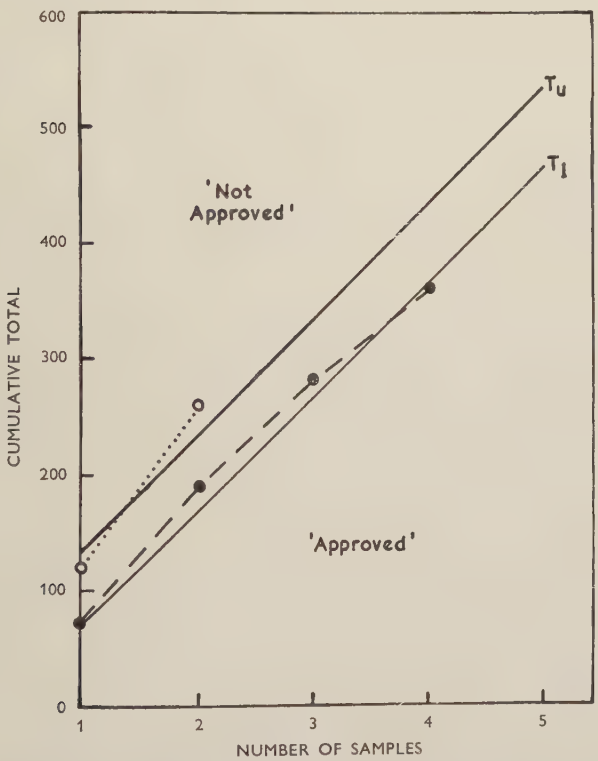


FIG. 1. Illustration of conventional sequential scheme. Broken line represents successive alues 70, 120, 90, 80—face 'approved'; dotted line represents successive values 120, 140—face 'not approved'.

The above procedure is simple and straightforward enough, provided that the distribution of variables follows the normal law. It breaks down, however, if the distribution is, say, lognormal. In this

case a similar procedure can be used by plotting the cumulative *product* of the results against the number of results; the corresponding limit lines then become curves. Alternatively the logarithm of the individual results can be taken, and a linear test can again be applied. In either case the essential simplicity of the original test is lost. In view of this an alternative test was sought which could be applied with equal simplicity when the variables were lognormally distributed.

Recommended Procedure

In order to obtain a test which can be applied for different types of distribution it is necessary to make it qualitative in form rather than quantitative, i.e. it is necessary to consider whether a result is greater or less than some critical value rather than to consider the numerical value of the result itself. It was decided to consider the following scheme:

'Let the upper and lower 95% points of the population whose median value is just on the critical value, μ, be λ_u and λ_l respectively (for a normal population $\lambda_u - \mu = \mu - \lambda_l = 2\sigma$); then a possible sequential scheme would be to reject the population if a single sample exceeded λ_u, or if two successive values lay between μ and λ_u. A similar criterion for acceptance would be adopted.'

It can be seen that the chance of a mistaken decision and the average sample number, for a population with true median m, depend only on the chance of individual sample values lying in the regions

$$< \lambda_l; \quad \lambda_l \text{ to } \mu; \quad \mu \text{ to } \lambda_u; \text{ and } > \lambda_u$$

The detailed calculations are given in the Appendix. Provided that the true distribution is easily transformed into the normal distribution, e.g. by the logarithmic transformation, the characteristics of the test are best obtained by carrying out the calculation for the normal model. The average sample number and the chance of making a mistake will be the same for corresponding percentage points in the real and normal distribution.

In order to obtain a more direct comparison with the conventional test we shall apply the recommended test to the example already discussed, i.e. for the normal distribution with $\mu = 100$, $\sigma = 20$, so that $\lambda_u = 140$ and $\lambda_l = 60$. In practical terms the test then becomes:

A working place will be '*approved*' if:

(a) Average conditions on a single shift are less than 60;

or (b) Average conditions on two successive visits both lie between 60 and 100.

A working place will be '*not approved*' if:

(a) Average conditions on a single shift are greater than 140;

or (b) Average conditions on two successive visits lie between 100 and 140.

The characteristics of such a sampling scheme can be seen from the llowing table:

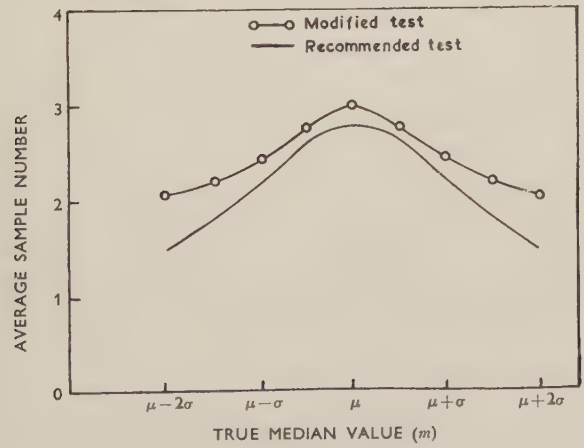
TABLE I

Operating characteristics of sampling scheme

True Median	Average Number of Shifts before Decision is taken	Maximum likely Number of Shift Samples required	Chance of a Wrong Decision
140	1·5	2	0·001
132	1·75	3	0·005
124	2·1	4	0·023
116	2·35	4	0·093
108	2·7	5	0·246
100	2·8	5	—
92	2·7	5	0·246
84	2·35	4	0·093
76	2·1	4	0·023
68	1·75	3	0·005
60	1·5	2	0·001

NOTE: The derivation of these figures is given in the Appendix.

It will be seen that the test fulfils the conditions required. The verage sample number is less than three, even for the most critical ase where $m = 100$, and it is most unlikely that more than five shifts ould have to be sampled before a decision was reached. Furthermore, ie chance of making a wrong decision when the population median



G. 2. Average sample number curves for the recommended test and the modified procedure.

differed from the critical value by σ is small (about 0·05) and the chance of a wrong decision when the difference is 2σ is negligible (0·001).

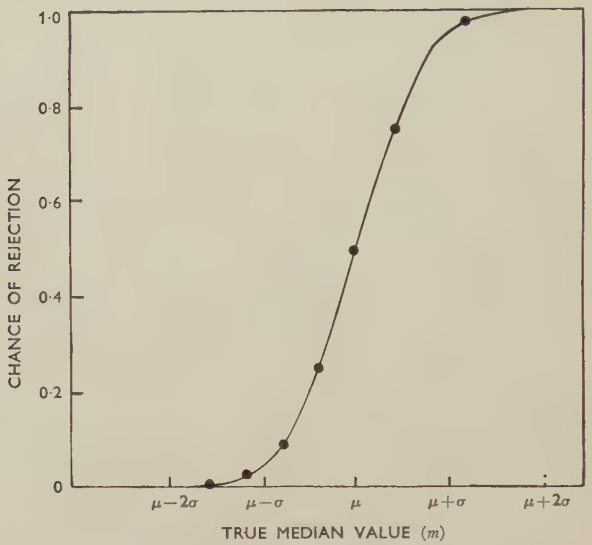


Fig. 3. Operating characteristic curve for the recommended test.

It is common practice to present the results quoted in the second and fourth columns of Table I in graphical form, and the appropriate curves are given in Fig. 2 (Average Sample Number) and Fig. 3 (Operating Characteristic). Methods of calculating these curves for the conventional test are well known, but they involve an approximation that breaks down for the small sample numbers considered here. Alternatively they could have been calculated by iteration, but this would be a lengthy process and I have not attempted it.

It may be instructive to give an example of the application of the recommended test when the variable being measured is lognormally distributed. Suppose that for a particular operation it has been found that the dustiness from shift to shift varied with the median dust concentration, the coefficient of variation being of the order of fifteen per cent. Suppose, furthermore, that in order to avoid contracting disease it is necessary to keep the average concentration over a long period below the level of 500 particles/c.c. In order to derive the test we need to know the 95% limits of the distribution when the median concentration is at the critical level, i.e. when $m = 500$. In logarithmic units the distribution will be normal with a median of $m^1 = 2{\cdot}699$ and a standard deviation σ given by

$$\sigma = \log_{10} 1{\cdot}15 = 0{\cdot}060\,7$$

The 95% limits of the distribution are then

$$2{\cdot}699 \pm 1{\cdot}96 \times 0{\cdot}060\,7, \quad \text{i.e. } 2{\cdot}580 \text{ and } 2{\cdot}818$$

Transformed back to the natural scale, the 95% limits become 380 and 658. The test may then be described as follows:

(*a*) The operation would be considered safe if either a single shift's sampling gave an average concentration below 380 particles/c.c. or if two successive shifts sampled had average concentrations between 380 and 500 particles/c.c.

(*b*) Similarly, the operation would not be considered safe if a single shift's sampling gave an average concentration above 658 particles/c.c. or if two successive shifts sampled had average concentrations between 500 and 658 particles/c.c.

Recommended Procedure when σ is not known

A simple modification of the recommended test makes it applicable when σ is not known. In this modification a face is accepted or rejected when two successive samples both have values less than or greater than the critical value μ. The characteristics of such a scheme can easily be calculated, and are given in the Appendix. The average sample number curve is given in Fig. 2 and as would be expected, rather larger sample numbers are involved; the operating characteristic so closely resembles that for the recommended procedure that the curve in Fig. 3 represents it with sufficient accuracy.

The modification of the conventional test to allow for an unknown standard deviation results in a non-linear sequential *t*-test. Several such tests have been described in the literature, but there are no simple methods of calculating the operating characteristics for such small sample numbers.

Reference may also be made to tests described by Barnard[2] and Armitage[3], where a value over the required level is given a positive score and that below the required level is given a negative score. The test is then carried out in the conventional manner, the total score being plotted against the number of samples. Such a test also avoids problems arising from non-normal distributions.

Comparison of the Recommended and Conventional Tests

Two sequential tests are usually compared on the basis of their operating characteristics and their average sample number curves. Although these have been given for the recommended test, they are not available for the conventional test, as has already been stated. Some attempts have been made to derive approximations to the curves for the conventional test, and these indicate that the operating characteristics and the average sample number curves are similar for the two tests. The conventional test must, from first principles, be the more efficient, but the indications are that the loss in efficiency using the recommended test is small.

A comparison of the relative merits of the two tests must therefore be based on the grounds of convenience and flexibility. The convenience

of the recommended test is clear; it involves no calculations, is simple in application however the variables are distributed (provided the distribution is unimodal), and could be operated by entirely untrained personnel. The drawback to the test lies in its lack of flexibility, owing to the fact that the test cannot easily be modified for other test conditions. The obvious way to modify the test is to use other lines than the 95% lines as the outer limits, but the fact that when the outer limits are discarded (in the procedure when σ is not known) the operating characteristic is virtually unaltered shows that this modification is not very effective. If we wished for a more stringent test, e.g. so that when

$m = \mu + \sigma$, $\alpha = 0{\cdot}001$ (instead of $\alpha = 0{\cdot}05$ as in the recommended procedure)

a new test would have to be found. With the conventional procedure such a modification is simply carried out. It would appear, therefore, that the recommended procedure has a limited field of application, but that within that field it has the advantage of simplicity of application.

Acknowledgement

This paper is published with the permission of the Director of Scientific Control, National Coal Board.

APPENDIX

Average Sample Number and Operating Characteristic of Recommended Procedure

For a given value m of the true median

Let p_1 be chance of a sample value exceeding λ_u
p_2 be chance of a sample value lying between λ_u and μ
p_3 be chance of a sample value lying between μ and λ_l
p_4 be chance of a sample value being less than λ_l

We can then construct the following table of chances:

n	*Chance of no decision after n samples*	*Chance of a decision at the n'th sample*
1	$p_2 + p_3$	$1 - (p_2 + p_3)$
2	$2p_2p_3$	$(p_2 + p_3) - 2p_2p_3$
3	$p_2p_3(p_2 + p_3)$	$p_2p_3[2 - (p_2 + p_3)]$
4	$2p_2^2p_3^2$	$p_2p_3[(p_2 + p_3) - 2p_2p_3]$
. . . . .		., etc.

The average sample number (A.S.N) is therefore:

$$[1 - (p_2 + p_3)] + 2[(p_2 + p_3) - 2p_2p_3] + 3[2p_2p_3 - p_2p_3(p_2 + p_3)] + 4[p_2p_3(p_2 + p_3) - 2p_2^2p_3^2] + \ldots$$

$$= 1 + 2p_2p_3 + 2p_2^2p_3^2 + 2p_2^3p_3^3 + (p_2 + p_3) + p_2p_3(p_2 + p_3) + \ldots$$

$$= \frac{2p_2p_3 + p_2 + p_3}{1 - p_2p_3} + 1 = \frac{(1 + p_2)(1 + p_3)}{(1 - p_2p_3)}$$

The chance of *accepting* a face

at the first sample is p_4
at the second sample is $p_4(p_2 + p_3) + p_3^2$
at the third sample is $p_4(2p_2p_3) + p_3^2p_2$
at the fourth sample is $p_4(p_2p_3)(p_2 + p_3) + p_3^2(p_2p_3)$
. . ., etc.

Thus the total chance, α, of *accepting* the face is given by

$$p_4[1 + (p_2 + p_3) + 2p_2p_3 + p_2p_3(p_2 + p_3) + 2p_2^2p_3^2 + \ldots] + p_3^2[1 + p_2 + p_2p_3 + p_2^2p_3 + p_2^2p_3^2 + \ldots]$$

whence $$\alpha = p_4\left\{\frac{p_2 + p_3 + 2p_2p_3}{1 - p_2p_3} + 1\right\} + p_3^2\left\{\frac{1 + p_2}{1 - p_2p_3}\right\}$$

or $$\alpha = \frac{1 + p_2}{1 - p_2p_3}\{p_3^2 + p_4(1 + p_3)\}$$

This relation gives the operating characteristic of the recommended procedure.

Characteristics of Procedure when σ is not known

The characteristics of the procedure when σ is not known can be obtained by putting

$$p_4 = 0,\ p_3 = p,\ p_2 = 1 - p$$

whence $$\text{A.S.N.} = \frac{(1 + p)(2 - p)}{1 - p + p^2}$$

and $$\alpha = \frac{(2 - p)p^2}{1 - p + p^2}$$

Preparation of Curves of Average Sample Number and Operating Characteristics

In Table I and in Figs. 2 and 3 it has been assumed that the variables are distributed normally with $\sigma = 20$. Thus for $m = 140$, $p_1 = 0{\cdot}5$, $p_2 = 0{\cdot}477$, $p_3 = 0{\cdot}023$, $p_4 = 0$, while for $m = 100$, $p_1 = p_4 = 0{\cdot}023$ and $p_2 = p_3 = 0{\cdot}477$. If the variables are distributed lognormally, the symmetrical form of Figs. 2 and 3 is only true for the logarithmically transformed variable. The curve for the real variables is not shown, but direct correspondence will be maintained between corresponding percentage points on the transformed and untransformed diagrams. Thus the upper 95% point will always have an A.S.N. of 1·5 and an α of 0·001.

REFERENCES

Davies, O. L., Editor (1954). *The Design and Analysis of Industrial Experiments*. Oliver and Boyd, Edinburgh and London.

Barnard, G. A. (1946). 'Sequential tests in industrial statistics', *Suppl. J. R. Statist. Soc.*, **8**, 1.

Armitage, P. (1947). 'Some sequential tests of Student's hypothesis', *Suppl. J. R. Statist. Soc.*, **9**, 250.

AN OPTICAL PERIODOGRAPH FOR USE IN THE ANALYSIS OF CONTINUOUS MEASUREMENTS

KENNETH W. KEMP
Research Department, British Nylon Spinners Limited

Dr Kemp describes an instrument which facilitates the analysis of records of a continuously varying quantity. The frequency of periodic components of variation may be measured and qualitative information may be obtained about components whose period is changing.

Introduction

Continuous records of the fluctuation of a constantly varying quantity are often met in the laboratories of modern industrial concerns. Records of this kind are visual and are usually obtained by the use of such devices as a cathode ray oscilloscope to which a photographic unit has been attached, or a pen recorder. An appropriate description of the type of data we shall consider would be 'continuous time series'. An example would be obtained if we were able to measure the cross-sectional area of a textile yarn at every point along its length, and to plot those measurements against distance along the yarn. Many instruments are available which do this automatically. Another example would be a record of the fluctuations in room temperature during a given period of time, if we could measure the temperature continuously.

The problem of analysing and interpreting these records, i.e. the problem of distinguishing systematic effects from random fluctuations, is not an easy one; in many instances the data are examined for their main features in a crude qualitative manner. The object of this paper is to bring an instrument which may be used to facilitate the analysis of continuous time series to the notice of statisticians who are concerned with them. The instrument will be referred to as the optical periodograph. No claim is made that the instrument necessarily obviates the need for a subsequent detailed analysis of the data. It is a useful preliminary tool which may be used to indicate what assumptions may be justified concerning the nature of the data. It will be found that the method is both speedy and accurate, an advantage which will appeal to the practical statistician.

The instrument to be described is basically very similar to one described by Douglass.[1,2] A development of this instrument, called the 'grating periodograph', which has been described by Foster,[3] has been used in assessing irregularities in cotton yarns. The Foster

ıstrument was specifically designed for harmonic analysis, and although xtremely useful for this purpose has thus a restricted field of applica- on.

Theoretical Considerations

The optical periodograph utilises the fact that when two line gratings ıounted on a transparent medium are superimposed one upon the ther a series of dark parallel bands or fringes appears when the lines f one grating are slightly inclined to those of the other. Such a set of ringes may be seen in Plate II(*a*). If the number of lines/inch on the rating (called the analysing grating) is known, then it is possible, by ıe use of a simple formula, to calculate from the geometry of the ringes the number of lines/inch on the other. To do this we need to now the angle between the gratings and the angle which the fringes ıake with the direction of the lines on the analysing grating. The ɔrmula is derived below. In Fig. 1:

θ = angle between the two gratings

ω = angle between the fringes and the analysing grating

s refers to the analysing grating

t refers to the second grating

s and t are the distances between the centres of two adjacent black lines

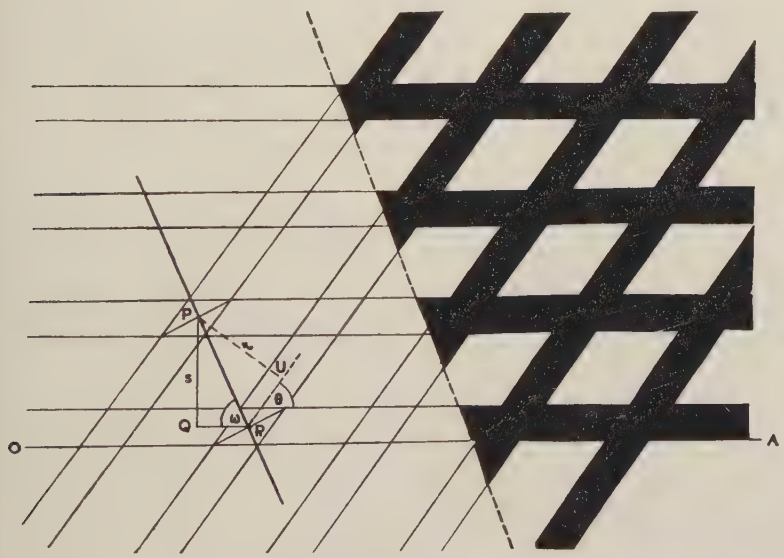


Fig. 1. Diagram showing the constants of gratings.

The line PR is the direction of the fringe and OA is a reference line. and R are the centres of the intersections of successive black lines on ıe two gratings, $\angle PQR = \angle PUR = \pi/2$. For clarity the opaque and

transparent areas of the two gratings are shown diagrammatically to the left of the dotted line.

$$PQ = s \qquad PU = t$$

From simple geometry:

$$\angle PRQ = \omega \quad \text{and} \quad \angle RPU = (\omega + \theta - \pi/2)$$

thus
$$PR = \frac{s}{\sin \omega} = \frac{t}{\cos (\omega + \theta - \pi/2)}$$

hence
$$t = s \cos \theta(1 + \cot \omega \tan \theta) \qquad \ldots.(1)$$

In Fig. 1 the dark lines are thinner than the spaces between them. They have been so drawn to clarify the diagram. In practice the best results are obtained when the spaces are thinner than the breadth of the dark lines (e.g. about three-quarters the breadth of the dark lines).

The appearance of parallel fringes when one grating is superimposed upon another is the criterion for the presence of a periodic pattern in the grating being analysed. Stated another way, it is the criterion that the lines on the latter grating are equally spaced. It is important to realise at this stage that if distances between the lines on the grating being analysed are not all equal but vary in a random manner, then a confused pattern will be seen.

The Instrument

The instrument is a simple optical device which converts the continuous time series into a pattern of parallel lines. Its main features are shown in Fig. 2. The transformation is achieved by means of a cylindrical lens.

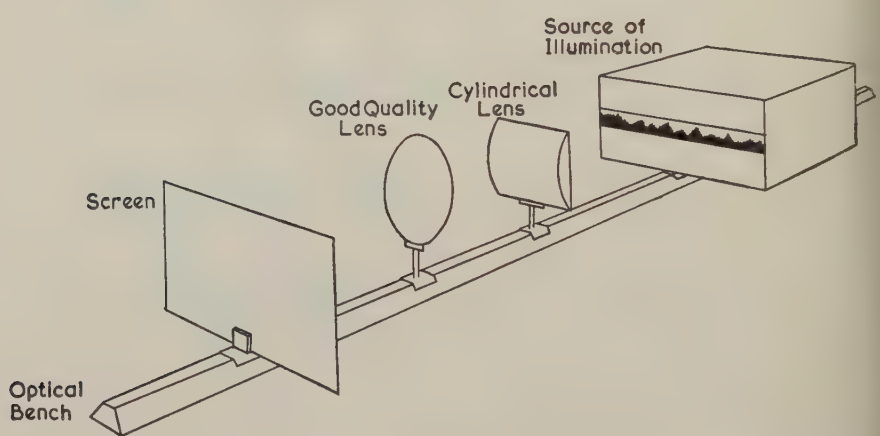


Fig. 2. The main features of the optical periodograph; the dimensions of the instrument constructed by the author were as follows:

(i) Box 18 in. wide × 18 in. long × 12 in. deep
(ii) Three 275-watt bulbs
(iii) Focal length of cylindrical lens = 3·10 in.
(iv) Focal length of convex lens = 5·25 in.

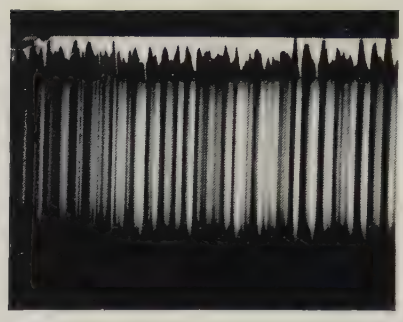
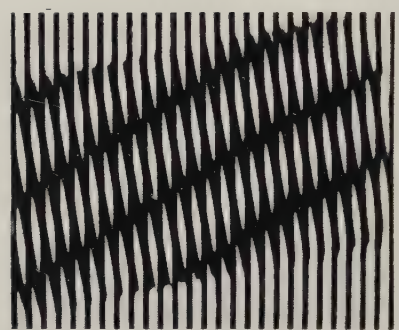
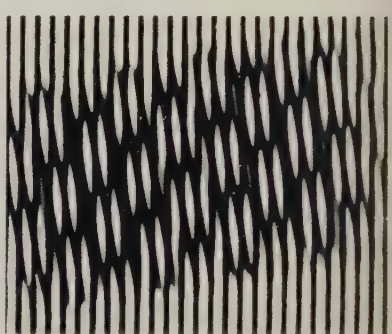


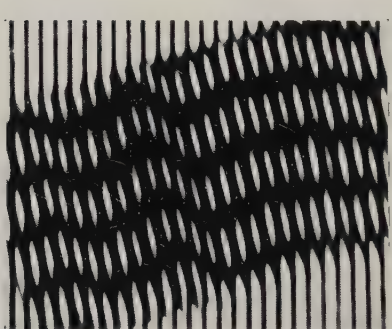
PLATE I. Silhouette and the corresponding line pattern.



(*a*)



(*b*)



(*c*)

(*d*)

PLATE II. Fringe patterns. (*a*) Single period, (*b*) double period, (*c*) varying period, (*d*) auto-regressive series.

A long, narrow strip is cut from the side of a box mounted on an optical bench, the strip being cut from the side facing down the bench. The box, which is used as a source of illumination, contains several powerful lamps. Two or three sheets of ground glass are placed between the lamps and the slit to ensure that the light is well diffused when it passes through the slit. A screen is mounted at the opposite end of the bench. In between the screen and the slit is placed a cylindrical lens with its cylindrical axis perpendicular to the direction of the optical bench, and parallel to and level with the line which bisects the narrow slit. A good photographic anastigmatic lens is placed between the cylindrical lens and the screen. If a silhouette of the continuous time series to be analysed is placed in front of the illuminated slit and the lenses adjusted so that a clear image is obtained on the screen, this image will be seen to consist of a number of parallel lines of varying density. An example is shown in Plate I. This effect is produced by the cylindrical lens, which distorts the image of the silhouette presented to it in a direction perpendicular to the lens axis but does not introduce any changes in a direction parallel to the axis. The continuous time series has thus been converted into a form which may be treated as a grating to be analysed as described previously. We may note in passing that if the series being analysed is strictly periodic the parallel lines on the screen are equally spaced. On the other hand, if the fluctuations in the *continuous* (*i.e. not discrete*) series are random in *time*, then nothing systematic is seen.

An analysing grating is now attached to the screen so that its lines are inclined to the vertical. A set of fringes similar to those shown in Plate II(*a*) will appear on the screen when the series being analysed contains a periodic component. Further, if the series being analysed consists of the sum of a periodic variable and a random component, the periodic component will produce a set of parallel fringes whereas the random component will show nothing. The instrument will therefore seek out the periodic component.

Interpretation of Patterns

The simplicity of the instrument and the principles involved enable one to predict, without ambiguity, the kind of fringe pattern which different types of variation will give. For example, we have seen that one periodic component gives rise to a single set of fringes. The slope of these is determined by the wavelength. If the series being examined contains two periodic components with different wavelengths there will be a set of parallel fringes for each period. A lattice pattern similar to the one shown in Plate II(*b*) is then obtained, since the two sets of fringes will have different slopes.

If the type of variation is oscillatory in form but the wavelength is changing smoothly the fringe pattern will not be a set of parallel straight lines, because the straight line is a consequence of the fixed period-length. The pattern in this case will be a set of equally spaced

smooth curves. An example obtained from a series of this kind is shown in Plate II(*c*).

If the series examined is an auto-regressive series of the form

$$u_x = au_{x-1} + bu_{x-2} + \epsilon_x$$

one would expect to obtain evidence of a fluctuating wavelength throughout the series examined, with nothing very systematic about the way in which this change takes place. Plate II(*d*) shows that this is so. The figure is a photograph of the pattern obtained on the instrument using one of Kendall's[4] auto-regressive series.

Evidently the instrument is capable of yielding useful information concerning the nature of the series under examination.

Sensitivity

Experience gained in use of the instrument soon reveals that the spacing of the analysing grating (i.e. the number of lines/inch) is of considerable importance. To achieve the best results with regard to the detection of a period and to attain the greatest accuracy it is essential that the number of lines/inch of the analysing grating shall be close to the number of lines/inch produced by the period being analysed. This is done by using a simple gearing system which joins the screen and the optical system. The gearing is arranged so that when the lenses are moved relative to the illuminated slit the screen moves in sympathy with them in such a way that the slit always remains in focus on the screen. The screen itself is constructed so that it acts as an analysing grating. It consists of a grating mounted with its lines inclined at an angle θ to the vertical. A horizontal line is drawn across the screen. This line is therefore inclined at an angle $\pi/2 - \theta$ to those of the grating. If a period is measured when its fringes are parallel to the line, then equation (1) transforms to

$$t = s \sec \theta$$

Hence the ratio t/s is fixed at a value which can be controlled by a suitable choice of θ. To estimate p, the wavelength in a series, we use the equation $p = \mathrm{r}(u).t$. Here $\mathrm{r}(u)$ is the value of the ratio (size of object)/(size of image) when the cylindrical lens is at a distance u from the silhouette. When $\mathrm{r}(u)$ is known (it can be obtained by experiment) it is obviously possible to calibrate the optical bench on which the instrument is mounted so that periods in a continuous time series can be measured directly from it. This method removes the possibility of overlooking a periodic component. The movement of fringes which results from varying the silhouette-to-lens distance gives an overall picture of the periodic components and their subharmonics. In practice it is possible to resolve a large number of periods. In the work for which I used the instrument the length of series examined was usually about 20 cm. With this length and a suitable choice of only two gratings it was found possible to cover a range of periods with

wavelengths from 0·20 cm. to 1·10 cm. It is thought that this range is adequate for most practical purposes.

Subharmonics

Experiment soon shows that it is easy to distinguish between fringes given by a periodic component and those due to one of its subharmonics. A component of wavelength λ has subharmonics of wavelength $\lambda.n$, where n is an integer. There is a characteristic disjointed pattern which one quickly learns to associate with these. The strength of the fringe pattern is much greater for the fundamental component.

The Silhouette

The preparation of silhouettes is not an arduous task, nor is it very time consuming. It is an advantage to record measurements on opaque paper, in which case there is no need to blacken the silhouette after it has been cut out.

Acknowledgements

It is a pleasure to acknowledge the help given to me in the construction of this instrument by Mr S. C. Simmens and Dr M. F. Culpin. I also wish to thank British Nylon Spinners for permission to publish this paper.

REFERENCES

[1] Douglass, A. E. (1914). 'Photographic periodogram of sun spot numbers', *Astrophysical Journal*, **40**, 326.
[2] Douglass, A. E. (1915). 'An optical periodograph', *Astro-physical Journal*, **41**, 173.
[3] Foster, G. A. R. (1946). 'Some instruments for the analysis of time series and their application to textile research', *J. R. Statist. Soc.*, B, **8**, 42.
[4] Kendall, M. G. (1949). 'Tables of auto-regressive series', *Biometrika*, **36**, 267.

CONFIDENCE LIMITS FOR THE RANGE OF A SEARCH RADAR

COLIN STEVENS

Communication from the Staff of the Applied Electronics Laboratories of the General Electric Company Ltd, Stanmore, England

The determination of confidence limits for the percentage points of a normal distribution is a problem with an exact, but somewhat troublesome, solution. In this article Mr Stevens describes a fairly simple approximate solution, which was found to be sufficiently accurate.

The function of a search radar is the detection of distant targets, and in general the greater the range of detection the better the radar may be said to be. Detection range is therefore an important aspect of search radar performance, and this paper is concerned with its determination from data obtained from trials.

There are, however, certain difficulties involved in the definition of radar range. In the first place it varies with the nature and aspect of the target, since a large target reflects more energy than a small one, and the reflecting area of a given target also varies considerably with its aspect. It is therefore customary to specify the nature and aspect of the target in some such phrase as 'Mosquito target, head-on'. Secondly, in the case of rapidly approaching targets, it has long been recognised that there is no such thing as 'the range' of a search radar, but rather a distribution of detection ranges. A number of causes contribute to this effect, including random variations in the strength of the received echo, the presence of unwanted and spurious random echoes on the display, variability of operators, and day-to-day variations in the setting up of the radar. It is essential to take this variability into account when defining detection range, and the definition in current use is framed in terms of the 'α% detection range' (α being some number between 0 and 100) defined as that range at or by which α% of incoming targets are detected. This definition is illustrated in Fig. 1, in which the upper figure shows the definition in terms of the probability density function (or frequency distribution) of detection ranges, and the lower in terms of the cumulative probability function. A numerical value of 75 is used for α in some radar specifications. Using x to denote the detection range (a random variable) and X_α to denote the α% detection range, the definition of X_α can be concisely expressed as

$$\Pr[x \geqslant X_\alpha] = \alpha \qquad (0 \leqslant \alpha \leqslant 1) \qquad \ldots.(1)$$

The problem with which this paper is concerned is the determination of the $\alpha\%$ detection range from experimental data. Specifically, given a random sample of N observed detection ranges, $x_1, x_2, \ldots, x_N$, what can be said about X_α? The natural approach is to plot the cumulative distribution of the data and thus make a direct estimate of X_α. This

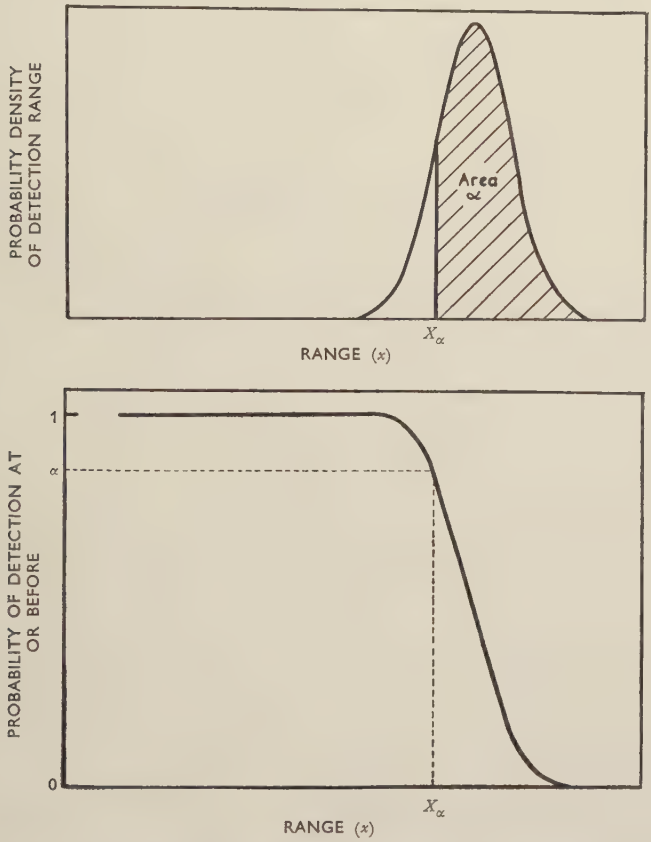


FIG. 1. Probability distribution of detection ranges; the upper diagram shows the probability density function and the lower diagram shows the cumulative probability function.

process is illustrated in Fig. 2, where a straight line has been fitted by eye to some hypothetical data plotted on normal probability paper; the data serve as the basis of a worked example discussed later in this paper. This graphical method has the advantage of dispensing with distributional assumptions. (A straight line has been fitted on normal probability paper in Fig. 2 because the data are known in this case to be normally distributed. The method may in principle be applied equally well to a smooth curve drawn through the same points plotted on ordinary squared paper.)

However, it is not of much use merely to estimate X_α without giving some indication of the degree of sampling fluctuation to which the estimate is liable. This fluctuation arises from the fact that X_α is

estimated from the data of the sample, and if the sampling process were repeated a number of times one would expect the samples to yield differing estimates of X_α.

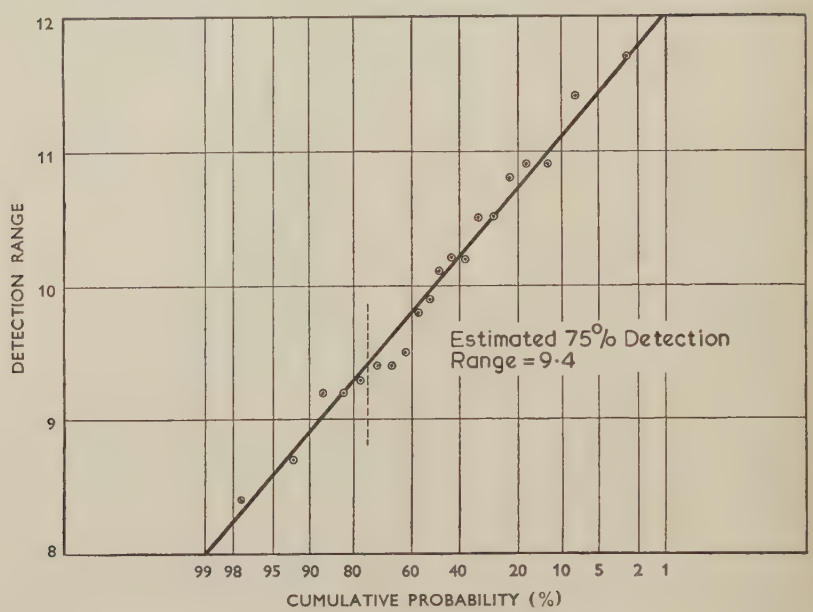


FIG. 2. Cumulative distribution of sample data on detection ranges plotted on normal probability graph paper.

A convenient way of expressing the precision of an estimate is by the statement of a confidence interval, defined by a lower limit X_1 and an upper limit X_2, in such a manner that the unknown quantity X_α can be asserted to lie between X_1 and X_2 with any desired probability γ that the assertion is correct. It should be remembered that X_α is not a random variable but an unknown constant of the distribution; the limits X_1 and X_2 are random variables. The statement really means that in long series of experiments a proportion γ of intervals (X_1, X_2) computed from each experiment would contain X_α and a proportion $1 - \gamma$ would not.

To determine X_1 and X_2 we need the sample mean $\bar{x}$ and standard deviation s defined by

$$\bar{x} = (x_1 + x_2 + \ldots + x_N)/N \qquad \ldots.(2)$$

$$(N - 1)s^2 = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_N - \bar{x})^2 \qquad \ldots.(3)$$

We also require K_ϵ, defined as the standardised normal deviate exceeded with probability ϵ. If t is normal with zero mean and unit variance, K_ϵ is defined by

$$\Pr[t \geqslant K_\epsilon] = \epsilon \qquad (0 \leqslant \epsilon \leqslant 1) \qquad \ldots.(4)$$

Values of K_ϵ may be obtained from standard tables of the normal integral: as examples, $K_{0.025} = 1{\cdot}960\,0$, $K_{0.25} = 0{\cdot}674\,5$, $K_{0.975} = -K_{0.025} = -1{\cdot}960\,0$.

The problem may be solved exactly by means of the non-central t distribution, and the method for doing this, with the necessary tables, is given by Johnson and Welch[1] (see Example 3 on page 374). The exact method, however, requires a troublesome iterative calculation, and interpolation in the tables. The present method, which is approximate, requires only the solution of a quadratic equation, and, for samples of size ten or more, yields solutions accurate enough for most practical purposes. A comparison of results obtained by the two methods is given in the Appendix for two examples, in each of which the sample size is ten. In neither case does the width of the confidence interval obtained by the approximate method differ by more than 2% from the exact value.

The basis of the approximation, that the variate $\bar{x} + ks$ is normally distributed with mean $\mu + k\sigma$ and variance $\sigma^2[1/\mathcal{N} + k^2/2(\mathcal{N} - 1)]$, is discussed in some detail by Johnson and Welch, and is shown to be a very good approximation for the present purpose.

The method now to be described is based on the assumption that the observations are independent random samples from a normal population of unknown mean and standard deviation. The limited amount of data available to me suggests that this is a reasonable assumption to make, and, from a theoretical point of view, the number of independent random processes affecting detection range (external and internal 'noise', operators, and variations in the tune of the set) could also be expected to lead to an approximately normal distribution of results.

The method is given in step-by-step form.

STEP 1. Compute $\bar{x}$ and s from the sample data.

STEP 2. Compute $\eta = (1 - \gamma)/2$ ....(5)

where γ is the probability to be attached to the confidence interval.

STEP 3. Obtain K_α and K_η from tables of the normal integral.

STEP 4. Solve the quadratic equation in k:

$$\left\{1 - \frac{K_\eta^2}{2(\mathcal{N} - 1)}\right\} k^2 - 2K_\alpha k + \left\{K_\alpha^2 - \frac{K_\eta^2}{\mathcal{N}}\right\} = 0 \qquad \ldots.(6)$$

Let the roots of this equation be k_1 and k_2, k_1 being the smaller.

STEP 5. The confidence limits X_1, X_2 are then given by

$$\left.\begin{aligned} X_1 &= \bar{x} + k_1 s \\ X_2 &= \bar{x} + k_2 s \end{aligned}\right\} \qquad \ldots.(7)$$

Numerical Example

The methods described will be illustrated by application to a hypothetical sample of detection ranges, obtained from a table of random

normal deviates with mean 10 and standard deviation 1. The sample data are as follows:

9·5	9·8	9·4	8·4	9·4
10·2	11·4	9·2	11·7	10·9
10·2	8·7	10·5	9·2	10·1
9·9	9·3	10·5	10·9	10·8

Suppose we are required to estimate the 75% detection range ($\alpha = 0{\cdot}75$). For the purpose of making a preliminary investigation the data have been plotted on normal probability paper in Fig. 2, by ordering the data in the form $x_1 \geqslant x_2 \geqslant \ldots \geqslant x_i \geqslant \ldots \geqslant x_N$, and plotting the ith ordered observation x_i at the position $[100(i - \frac{1}{2})/N]\%$. The use of these plotting positions is discussed by Chernhoff and Liebermann,[2] who show that they yield a quite efficient estimate of the population standard deviation, and hence presumably an efficient estimate of a percentage point of the distribution. By means of a line fitted by eye to the data, the 75% detection range can be estimated as 9·4. This happens to be in good agreement with the true 75% detection range of the population from which the sample was drawn, i.e. $10{\cdot}00 - 0{\cdot}67 = 9{\cdot}33$.

Suppose we are now required to determine the 95% confidence interval for the 75% detection range, i.e. $\gamma = 0{\cdot}95$.

STEP 1 $\bar{x} = 10{\cdot}00$
$s = 0{\cdot}880\,8$

(That $\bar{x}$ is identical with the population mean is pure coincidence.)

STEP 2 $\eta = (1 - 0{\cdot}95)/2 = 0{\cdot}025$

STEP 3 $K_\alpha = K_{0{\cdot}75} = -K_{0{\cdot}25} = -0{\cdot}674\,5$
$K_\eta = K_{0{\cdot}025} = 1{\cdot}960\,0$

STEP 4. Since in this case $N = 20$, the quadratic equation in k becomes

$$(0{\cdot}898\,91)k^2 + 2(0{\cdot}674\,5)k + 0{\cdot}262\,87 = 0$$

yielding the roots

$$k_1 = -1{\cdot}271$$
$$k_2 = -0{\cdot}230$$

STEP 5. The required confidence limits are therefore

$$X_1 = 10{\cdot}00 - (1{\cdot}271)(0{\cdot}880\,8) = 8{\cdot}88$$
$$X_2 = 10{\cdot}00 - (0{\cdot}230)(0{\cdot}880\,8) = 9{\cdot}80$$

We therefore make the assertion, with a 95% chance of being right, that the 75% detection range of the radar lies between 8·88 and 9·80.

Number of Observations for an Interval of Given Width

The result of (6) may be used to determine the width W of the confidence interval, defined by

$$W = |X_1 - X_2| \qquad \ldots\ldots(8)$$

From (7): $$W = |k_1 - k_2|s \qquad \ldots\ldots(9)$$

By solving (6), or otherwise, we can obtain an expression for $|k_1 - k_2|$, and (9) may then be written

$$\frac{W}{s} = 2K_\eta \frac{\left\{\frac{1}{N} + \frac{K_\alpha^2}{2(N-1)} - \frac{K_\eta^2}{2N(N-1)}\right\}^{\frac{1}{2}}}{\left\{1 - \frac{K_\eta^2}{2(N-1)}\right\}} \quad \ldots\ldots(10)$$

This expression is valuable as a guide to the number of observations required to establish a confidence interval of predetermined width (e.g. 10 miles) with any given degree of confidence γ. This can only be done, however, if a reliable estimate of σ is available beforehand which can be used for s in (10). Another feature of this result is that, for large N, the width W varies inversely as the square root of N, so that, for example, four times the number of observations are required to reduce the width of the interval to a half.

The ratio W/s has been plotted in Fig. 3 for 95% confidence intervals for three different values of α, 0·75, 0·95, 0·99.

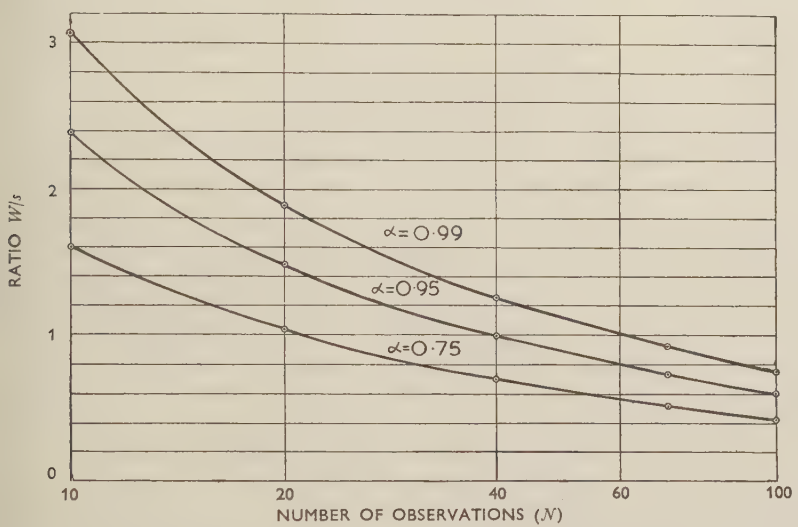


FIG. 3. Variation of width of 95% confidence limits for X_α with number of observations. W = width of confidence interval; s = sample standard deviation; $\gamma = 0{\cdot}95$.

Example

Suppose we have to estimate beforehand the number of observations required to determine the 75% detection range to an accuracy of 1 unit with 95% confidence. Previous experience suggests a value for σ of ·7 units. Then $W = 1$, and we take $W/s \simeq 1/1{\cdot}7 = 0{\cdot}59$. Reading across the figure on to the $\alpha = 0{\cdot}75$ curve, we obtain $N = 61$. This then is our estimate of the number of observations required.

In estimating W by these means it should be remembered that its accuracy depends on the accuracy of the original estimate of σ, and also that σ has been used for s in the calculation, thereby making no allowance for fluctuations in the sample standard deviation s.

APPENDIX

Let x denote detection range, and let X_α denote the range such that

$$\Pr[x \geqslant X_\alpha] = \alpha \qquad \ldots.(A\,1)$$

The problem is to determine limits X_1, X_2 $(X_1 < X_2)$, computed from a sample of N observations in such a manner that the interval (X_1, X_2) can be asserted to contain the point X_α with probability γ that the assertion is correct. Since there is an infinite number of such intervals, we shall add the further restriction that we seek a central confidence interval, i.e. one satisfying the following equations:

$$\Pr[(-\infty, X_1) \text{ contains } X_\alpha] = (1-\gamma)/2 = \eta \qquad \ldots.(A\,2)$$

$$\Pr[(X_2, \infty) \text{ contains } X_\alpha] = (1-\gamma)/2 = \eta \qquad \ldots.(A\,3)$$

in which $(-\infty, X_1)$ denotes the interval with $-\infty$ and X_1 as its end points, etc.

It is assumed that the variate x is normally distributed with unknown mean μ and unknown standard deviation σ. In consequence of this we have

$$X_\alpha = \mu + K_\alpha\sigma \qquad \ldots.(A\,4)$$

This leads us to investigate functions of the type

$$X_i = \bar{x} + k_i s \qquad (i = 1, 2) \qquad \ldots.(A\,5)$$

for the limits X_1, X_2. The X_i are statistical variates, being functions of the sample values, and it is therefore legitimate to write (A 2), (A 3) as

$$\Pr[X_1 \geqslant X_\alpha] = \eta \qquad \ldots.(A\,6)$$

$$\Pr[X_2 \leqslant X_\alpha] = \eta \qquad \ldots.(A\,7)$$

The first of these can be written in the form

$$\Pr[\bar{x} + k_1 s \geqslant \mu + K_\alpha\sigma] = \eta$$

which can be rearranged as follows:

$$\Pr\left\{\frac{\sqrt{(N)}(\bar{x}-\mu)/\sigma - K_\alpha\sqrt{N}}{s/\sigma} \geqslant -k_1\sqrt{N}\right\} = \eta \qquad \ldots.(A\,8)$$

This establishes the connection with non-central t, since this is defined as

$$t = (z + \delta)/\sqrt{\omega} \qquad \ldots.(A\,9)$$

in which z is normal with zero mean and unit variance, and ω is independently distributed as χ^2/f, f being the degrees of freedom of the χ^2. The point $\mathrm{t}(f, \delta, \epsilon)$ is defined as the number such that

$$\Pr[t > \mathrm{t}(f, \delta, \epsilon) | f, \delta] = \epsilon$$

and it follows therefore that

$$-k_1\sqrt{N} = \mathrm{t}(N-1, -K_\alpha\sqrt{N}, \eta) \qquad \ldots.(A\,10)$$

The tables of non-central t permit this equation to be solved exactly for k_1, and k_2 can be found in a similar manner. In practice, however, this can be a troublesome process, and we shall therefore seek an approximate solution for k_1 and k_2 which can be found in a simple manner.

Approximate Solution for Confidence Limits

Consider $X_1 = \bar{x} + k_1 s$ as a linear combination of two variates $\bar{x}$ and s. Under the normality assumed for the observations the following remarks may be made:

(i) $\bar{x}$ and s are independent.

(ii) $\bar{x}$ is normally distributed with mean μ and variance $\sigma^2/\mathcal{N}$.

(iii) The distribution of s tends to normality with increasing $\mathcal{N}$, with mean and variance given by

$$\mathrm{E}(s) = a\sigma \qquad \ldots.(\mathrm{A}\ 11)$$

$$\mathrm{Var}(s) = b^2\sigma^2/2(\mathcal{N} - 1) \qquad \ldots.(\mathrm{A}\ 12)$$

The constants a and b depend on $\mathcal{N}$, and approach unity as $\mathcal{N}$ increases. For example, when $\mathcal{N} = 10$, $a = 0{\cdot}973$, $b = 0{\cdot}985$. *We shall hereafter assume that* $a = b = 1$, an assumption discussed by Johnson and Welch[1] and shown to lead to a more accurate solution of the present problem than that obtained by giving them their correct values.

(iv) Since $\bar{x}$ is normal and s nearly normal, the linear combination $\bar{x} + ks$ is also nearly normal, and more so than s.

We shall therefore assume X_i to be normal, with mean and variance given by

$$\mathrm{E}(X_i) = \mu + k_i\sigma \qquad \ldots.(\mathrm{A}\ 13)$$

$$\mathrm{Var}(X_i) = \sigma^2/\mathcal{N} + k_i^2\sigma^2/2(\mathcal{N} - 1) \qquad \ldots.(\mathrm{A}\ 14)$$

Now consider (A 6):

$$\Pr[X_1 \geqslant X_\alpha] = \Pr\left\{\frac{X_1 - \mathrm{E}(X_1)}{\sigma_{X_1}} \geqslant \frac{X_\alpha - \mathrm{E}(X_1)}{\sigma_{X_1}}\right\} = \eta$$

and therefore

$$X_\alpha - \mathrm{E}(X_1) = K_\eta\sigma_{X_1} \qquad \ldots.(\mathrm{A}\ 15)$$

From (A 4), (A 13), (A 14) we have

$$K_\alpha - k_1 = K_\eta[1/\mathcal{N} + k_1^2/2(\mathcal{N} - 1)]^{\frac{1}{2}} \qquad \ldots.(\mathrm{A}\ 16)$$

A similar argument applied to (A 7) yields, for k_2:

$$k_2 - K_\alpha = K_\eta[1/\mathcal{N} + k_2^2/2(\mathcal{N} - 1)]^{\frac{1}{2}} \qquad \ldots.(\mathrm{A}\ 17)$$

By squaring both sides of (A 16), (A 17) we see that k_1 and k_2 are roots of a quadratic equation in k:

$$[1 - K_\eta^2/2(N-1)]k^2 - 2K_\alpha k + [K_\alpha^2 - K_\eta^2/N] = 0 \quad \ldots.(\text{A } 18)$$

which is the required result.

A Note on the Accuracy of the Approximation

In order to evaluate the accuracy of the approximation, exact and approximate solutions of two examples have been computed for comparison.

CASE I $N = 10, \alpha = 0{\cdot}75, \gamma = 0{\cdot}95$

CASE II $N = 10, \alpha = 0{\cdot}10, \gamma = 0{\cdot}90$

(This is Example 3 in Johnson and Welch's paper.[1])

The exact and approximate solutions are as follows:

CASE I

Exact	*Approximate*
$k_1 = -1{\cdot}673$	$k_1 = -1{\cdot}661$
$k_2 = -0{\cdot}056$	$k_2 = -0{\cdot}054$
$\lvert k_1 - k_2 \rvert = 1{\cdot}617$	$\lvert k_1 - k_2 \rvert = 1{\cdot}607$

CASE II

Exact	*Approximate*
$k_1 = 0{\cdot}712$	$k_1 = 0{\cdot}696$
$k_2 = 2{\cdot}355$	$k_2 = 2{\cdot}321$
$\lvert k_1 - k_2 \rvert = 1{\cdot}643$	$\lvert k_1 - k_2 \rvert = 1{\cdot}625$

REFERENCES

[1] JOHNSON, N. L., and WELCH, B. L. (1940). 'Applications of the non-central t distribution', *Biometrika*, **31**, 362.

[2] CHERNHOFF, H., and LIEBERMANN, G. J. (1954). 'Use of normal probability paper', *J. Amer. Statist. Ass.*, **49**, 778.

NOTES AND COMMENTS

TATISTICIANS will have been pleased to learn that in the last Birthday Honours List a knighthood was awarded to Mr H. Campion, CB, CBE, Director of the Central Statistical Office, Cabinet Office. Sir Harry Campion is President of the Royal Statistical Society for the current session.

★ ★ ★

The results of a small survey by students attending a three weeks' full-time course on Statistical Quality Control and Acceptance Sampling at the College of Technology, Birmingham, should be of particular interest to some of our readers. They were published in the March–April issue of *The Inspection Engineer incorporating Quality Control* in an article entitled 'Quality Control in the Midlands' by C. J. Anson.

The survey was divided into four sections, dealing in turn with acceptance sampling procedures, quality control procedures, the division of responsibilities for quality control, and the economics of quality control. The pattern emerging from the results of the survey shows an inspection force using statistically based sampling schemes and statistical quality control, and plotting control charts. The answers to the questions on responsibilities and procedures were clear and definite. On the other hand, the answers to the questions on economic aspects of quality control suggest that there was not the same clarity in the evaluation of the costs involved, either the savings achieved or the cost of achieving them. Perhaps the most useful part of the article is the questionnaire itself, which could form a very useful catechism for statistical quality controllers.

★ ★ ★

It is now possible for suitably qualified persons in the London area to do part-time or full-time work at Birkbeck College for the degrees of Ph.D. and M.Sc. in Statistics. Further particulars and forms of application may be obtained from the Registrar, Birkbeck College, Malet Street, London,W.C.1.

★ ★ ★

At the invitation of the City and University of Edinburgh and the Royal Society of London, the International Congress of Mathematicians will meet in Edinburgh from 14th to 21st August 1958. His Royal Highness the Duke of Edinburgh has graciously consented to extend his patronage to the Congress. Among the sections into which the subject matter is divided is one on Probability and Statistics. Further information may be obtained from the Secretary of the Congress, Dr Frank Smithies, Mathematical Institute, 16 Chambers Street, Edinburgh, 1.

MEETINGS OF SECTIONS OF THE ROYAL STATISTICAL SOCIETY

INDUSTRIAL APPLICATIONS SECTION

BIRMINGHAM AND DISTRICT GROUP

3rd April 1957. Systems of Stock Control: F. I. Musk

15th May 1957. One-day Conference: Statistics and Stock Control. A Survey of Stock Control Problems: J. Harling. 'Models' of Stock Control Systems: A. M. Lee. Probability and Stock Control: R. R. P. Jackson. Inventory Control and Lot Size in Batch Production: A. H. Schaafsma.

LONDON GROUP

5th April 1957. Symposium on Electronic Computers. Economic Aspects: M. Hancock. Methodology: L. C. Payne. Recent Advances: S. L. H. Clarke

3rd May 1957. Some Simple Experiments: D. R. Cox

MERSEYSIDE GROUP

2nd April 1957. Marginal Cost: A. Muir

SHEFFIELD GROUP

25th April 1957. Machine Translation and other Uses of Electronic Computers: A. D. Booth

STUDY SECTION

LONDON GROUP

21st March 1957. Can the Statistician help the Investor? P. Freeman

24th April 1957. The Pattern of Consumer Purchases: A. S. C. Ehrenberg

8th May 1957. A Discussion of Some Problems of Statistical Inference: G. A. Barnard, M. G. Kendall, B. L. Welch

BRISTOL GROUP

21st March 1957. National Income Studies in the Colonies: A. R. Prest

Dr Prest discussed the problems encountered in making national income investigations in under-developed countries, giving examples of the difficulties in defining productive activity, of the evaluation of production, and of western classifications rendered inappropriate by native customs. He then spoke of progress since 1945, of the usefulness of the information which has become available, and of the forms of future developments.

MEDICAL SECTION

28th May 1957. Some Aspects of Cigarette Poisoning: J. H. Edwards

BOOK REVIEWS

Arthur O. Dahlberg: *National Income Visualized*. New York: Columbia University Press; London: Oxford University Press, 1956. Pp. x + 117. 28s.

Dr Dahlberg, of the Visual Economics Laboratory of Columbia University, has produced an impressive exposition of the elements of national income statistics and national accounting. The main feature of the book is a series of fifty flow charts, illustrating the streams of money and of goods and services which are measured in national accounting. The flow charts are accompanied by a very clear explanatory text. The series is progressive, beginning with the simple economic relationship between the worker-consumer and the productive enterprise, and gradually introducing the economic and statistical complications involved in the operations of government, savings and investment, and foreign trade. We begin, as it were, with the single circular track and two stations of the child's first model railway and end with the full apparatus of up lines and down lines, points, level crossings, tunnels, and sidings that fill us with wonder at the model railway exhibitions.

The system of accounting illustrated is, of course, that used by the US Department of Commerce. Only minor adaptations would be required in using the book to expound the British system. The most important and very necessary change would be to introduce foreign trade somewhat earlier than Diagram 47. The book also contains a complete set of forty-one tables from the Department of Commerce's National Income Supplement of 1954 (giving figures from 1929 to 1953)—a volume which corresponds to the British National Income *Blue Books*.

Flow charts have, of course, been used before to illustrate the system of national accounting [for example, the British system is exhibited in this way by D. K. Burdett ('Social accounting in relation to economic theory', *Econ. J.*, **64**, 679)], but never, to my knowledge, in so elaborate a way. Moreover, the diagrams are more elaborate in construction than the two-dimensional flow charts. By perspective drawings the artist is able to use two or more 'storeys' in his diagram. This serves a valuable expository purpose, since it makes it possible to introduce pictorially stocks as well as flows. Thus receipts from sales do not enter one end of the schematic factory and emerge immediately as wage payments, etc., at the other. The sales receipts are (more realistically) routed to a lower storey representing the factory's bank account (or stock of liquid assets), from which a series of payment flows emerges. Thus expenditure can be shown visually to depend on the assets position of the spender—whether enterprise or household—as well as on his income. This point can easily be lost sight of, not only in the exposition of national accounting but also in the analysis of economic affairs. Dr Dahlberg is insistent on the importance of supplementing national accounts by statistics of assets and liabilities.

In other respects, too, Dr Dahlberg's exposition goes beyond the national accounts. Thus he introduces the basic concepts of input–output tables and makes clear their relationship with the national accounts. He also gilds the

pedagogic pill by a rather tantalising summary of his own short-term forecasts of national product. In spite of its sometimes too enthusiastic tone, the book is a most helpful guide to the beginner who wishes to understand national accounting, and to the teacher who is willing to make use of visual methods.

It is unfortunate that the author is unable to devote any space to the sources from which the statistics are drawn. Some basic knowledge of the kinds of raw material from which highly processed statistics are produced seems an essential part of the equipment of the intelligent user. This is not because the administrative machinery for collecting statistical data is important in itself to others than the form-fillers and the statisticians; but anyone who uses economic statistics seriously should be able to assess their probable accuracy and weaknesses from at least a general knowledge of the ways in which they are compiled. C. T. SAUNDERS

Sir Ronald A. Fisher: *Statistical Methods and Scientific Inference*. Edinburgh and London: Oliver and Boyd, 1956. Pp. viii + 175. 16s.

Sir Ronald Fisher's new book follows hard on his recent article on this subject in the *Journal* (*Series B*), and will be avidly scrutinised by all who have been following the exchange of views which this article provoked.

Following a somewhat truculent foreword, a historical review of the development of the ideas of probability in relation to scientific inference leads to a fundamental chapter on the rationale of tests of significance and fiducial inference. This is followed by a chapter on some controversial tests of significance, including the two-by-two table and the Behrens test. Next come some specific examples of inferences involving probability and likelihood, and the book ends with a chapter on the principles of estimation. The whole book contains plenty of food for thought and argument, and plenty of controversial material stripped to essentials.

Those of us who teach (for whom this book was partly written) cannot escape the duty of trying to understand Fisher's fiducial argument and explain it to our students; we are always on the look-out for a clear and authoritative statement of this argument expressed, if possible, in watertight logical form. Here we are somewhat disappointed; the argument in Chapter III leans rather uncomfortably on some special pleading involving the exponential distribution, and in one important place is perplexingly ambiguous.

As to the first criticism, the exponential distribution, for which

$$\mathrm{d}f = \theta \mathrm{e}^{-\theta x}\mathrm{d}x \qquad (\theta > 0,\ 0 \leqslant x < \infty)$$

is such that the sum X of n random values is distributed as $\chi^2_{2n}/2\theta$; hence the probability statement

$$\mathrm{prob}(\theta \leqslant k/2X) = \mathrm{prob}(\chi^2_{2n} \leqslant k)$$

is not only a statement of fiducial probability but also unfortunately happens to be true when interpreted as a long-run relative frequency in a sequence of sums of n random drawings from any arbitrary sequence of exponential distributions. As to the second, a crucial step in Sir Ronald's explanation of the fiducial argument hinges on a reference set obtained *from all values of* θ, a phrase which is not clarified. Does it mean some population of values of θ, or the mathematical interval $0 \leqslant \theta \leqslant \infty$? If the first, the population, and therefore the reference set, is not defined; if the latter, whatever it is that is obtained is surely not a reference set at all.

This is definitely a book for the mature statistician, and will undoubtedly join its predecessors, *Statistical Methods for Research Workers* and *The Design of Experiments*, as a classic in the literature of statistics. Although it cannot be recommended as a textbook for students, they will find it provocative and interesting reading, and will doubtless be amazed at the intensity of the emotions which can be aroused in arguments on statistical fundamentals.

G. H. JOWETT

Maurice G. Kendall and William R. Buckland (prepared by): *A Dictionary of Statistical Terms*. Published for the International Statistical Institute. Edinburgh and London: Oliver and Boyd, 1957. Pp. xi + 493. 25s.

In 1951 the International Statistical Institute commissioned Professor Kendall to supervise the preparation of a dictionary of statistical terms. The dictionary has been compiled in Professor Kendall's department in the London School of Economics, and it has now been published in an elegant and inexpensive volume.

The dictionary itself occupies two-thirds of the space, the remainder being devoted to glossaries of French, German, Italian and Spanish terms with their English equivalents. The inclusion of these glossaries greatly increases the usefulness of the book; indeed, it is a pity that it was not possible to include also glossaries of the foreign equivalents of English terms. To find acceptable German for *quasi-random sampling* is probably more difficult than to discover from an ordinary German-English dictionary the meaning of *zufallsähnliches Stichprobenverfahren*.

Professor Kendall has described (*Bull. Inst. Statist.*, 1954, **34**, part 2, 629) some of the problems encountered in constructing the dictionary proper and the methods adopted in dealing with them. It was decided to exclude most purely mathematical terms—even those commonly used in statistics, since adequate mathematical dictionaries already exist—and terms which represent applications, in particular fields, of general statistical ideas. Although the omission of some borderline cases may be regretted, I agree in general with the authors in their choice of terms to be included. Nevertheless I cannot feel that the dictionary, taken as a whole, is satisfactory. It is difficult to imagine what kind of user would find it valuable. Two classes of potential users come to mind at once: statisticians seeking for an explanation of a term used in a branch of statistics which is unfamiliar to them, and laymen trying to understand a statistical report. It is to the latter particularly that the dictionary might be a real asset; statisticians should know where to find, in textbooks and papers, the explanation they seek. Technical dictionaries are not generally used by the technically qualified, and have little or no place in the instruction of the student. The style of the dictionary is, however, particularly unsuitable for the layman, for whom it is too mathematical and rigorous. For example, a *Stochastic Process* is defined as 'A family of variates $\{x_t\}$ where t assumes values in a certain range T. In most practical cases x_t is the observation at time t and T is a time-range but t may also refer to distribution in space and may be considered for discontinuous or continuous values . . .'. The section continues with mathematical definitions of stochastic continuity, differentiability, and integrability. Under *Stochastic* itself a far better dictionary definition is given '. . . a stochastic process is one wherein the system incorporates an element of randomness as opposed to a deterministic system'.

It was not to be expected, with such authors, that any actual errors would be found in the definitions. Some entries are difficult to understand, and many seem too terse. In general there is too little explanation of the reason for the introduction of a statistical term: *Comparative Mortality Figure* is defined rather clumsily, in terms of standardised and crude death rates (neither of which is defined) and the Laspeyres' form of index-number, but no indication is given that its purpose is to make comparisons possible between populations of different age distribution. Yet occasionally explanations are included; under *Fertility Gradient* it is remarked that one aim of randomisation and other devices of experimental design is to eliminate bias due to the existence of such gradients. Strictly, this remark has no place in the definition, but it enormously adds to its value. It is a pity that abbreviations have been excluded. To take an example already used, most readers needing a definition of *Comparative Mortality Figure* are likely to have met it only as *C.M.F.* The abbreviations *S.D.* and *S.E.* are also often found without explanation, regrettably enough, in column headings of numerical tables in published papers. At least one definition is tautological; under *Fiducial Distribution* we find 'A distribution of a parameter required for fiducial inference (q.v.) . . .' and under *Fiducial Inference:* 'A type of statistical inference based on fiducial distribution (q.v.) . . .', and the remainder of these paragraphs does little but enlarge the circle. Lexicographers are perhaps allowed at least one rather malicious joke.

Criticisms such as these amount to little when set against the vast number of entries which provide, in the particular style adopted, a clear and accurate definition. Particularly valuable is the inclusion of definitions of terms introduced and used by the Italian school of statisticians. In summary, there is no doubt that the authors have made a valiant attempt at a laborious and almost impossible task, but in my opinion they have not been entirely successful because they did not sufficiently consider who was likely to use their work before choosing the manner of presentation. The result is a volume which is not very suitable even for the rather limited group of readers who might be thought to require it. P. D. OLDHAM

J. Aitchison and J. A. C. Brown: *The Lognormal Distribution, with special reference to its uses in economics.* University of Cambridge Department of Applied Economics Monograph No. 5. Cambridge University Press, 1957. Pp. xviii + 176. 35s.

The study of the distribution of a variate of which the logarithm is normally distributed has a long history in statistics; the work done, however, has been rather spasmodic and its publication has been widely scattered. The authors are to be commended on the skill with which they have collected together and summarised in a systematic way the previous work, both on the theory of the lognormal distribution and on its practical applications. In addition, they have made useful and interesting contributions of their own on both the theory and the applications.

After a brief history of the lognormal distribution the authors give a formal development of the main properties of the distribution. Since many of the theorems stated follow almost trivially from the properties of the normal distribution, proofs have wisely been omitted; the few proofs that are given seem to be chosen more because the authors wish to use the theorems in their own applications than for their difficulty or for the more general importance of the properties concerned.

The book justifiably contains a long discussion of problems of estimation, since the computation involved in efficient methods of estimation is frequently heavy, yet (as the authors point out) doubts about the efficiency of the simpler methods may have caused many to avoid the use of the distribution. Various methods are discussed both for estimating the mean and variance of the lognormal distribution starting at zero and for estimating the parameters of the corresponding normal distribution; in addition, methods of estimation for the three-parameter lognormal distribution with an unknown non-zero starting point and for truncated and censored distributions are discussed. The problems of estimation are tackled by a theoretical approach and by the use of calculations made on the basis of sixty-five artificially constructed samples of various sizes from lognormal populations with a number of different values for the standard deviation of the corresponding normal distribution. This second approach, which can help to illustrate the points brought out in the theoretical discussion, is particularly useful where, as here, one of the methods considered is a graphical one the efficiency of which it would not otherwise be possible to assess. The authors elaborate on this use of artificial samples by working with separate graphical estimates made by persons with different experience.

Unfortunately the use of artificial samples in the book may be a little marred by the fact that the samples chosen do not come very satisfactorily through tests for lognormality. The authors apply tests of skewness and kurtosis for a two-parameter distribution to their samples. In one case eight of the sixty-five samples gave significant results at the five per cent level; in the other, where different tests were used for different-sized samples, ten of the sixty-five samples gave significant results at the same level. Both of these figures seem rather high—even the former is itself significant at the two per cent level. The different methods of estimation are compared on the basis of the root-mean-square of the deviations of the estimated from the true values; this is done first grouping together all samples of the same size and then grouping together samples by ranges of the standard deviation of the normal population. If one compares the squares of the figures given for estimates made by the method of maximum likelihood of, for example, the mean of the normal distribution with the corresponding expected values, one finds in several cases considerable differences; these presumably result from the nature of the samples used. It may be true that the comparison of the efficiencies of different methods is not seriously impaired by the unsatisfactory nature of the samples, but one would like further reassurance on this before accepting the results.

One chapter is devoted to probit analysis; in addition to presenting methods of estimation, the authors contribute an interesting section on problems of convergence in quantal response computations. In their later discussion of the analysis of consumers' behaviour they make further use of the techniques of probit analysis in their attempts to present a new model for such analysis. To an economist the chapter in which these attempts are made must remain one of the most stimulating parts of the book, even if one eventually decides that they are not completely successful. The initial assumption that the consumer's preferences for a single commodity with a given set of prices can be expressed as a number of points of his income at which he would buy an extra unit until a point of saturation is reached, justifies the claim to be weak; it seems both reasonable and attractive.

Doubts, however, might be expressed as to whether the assumptions involved in the processes of aggregation and in allowing for price changes are so reasonable. The former involves, for example, the assumptions that the frequency of occurrence of points of purchase of extra units at different income levels follows the lognormal distribution for the whole population and that the people with a given income at a family budget enquiry can be taken as representing the behaviour of the whole population if they had that income. Neither assumption is verifiable, and the only test is the goodness of fit of family budget data to the implied Engel curve, which, with available data, is not a very strong test.

The introduction of prices into the analysis does not appear very elegant and involves the rather dubious use of an index number of prices. There also seem to be two mistakes at this stage. The elasticity of substitution is confused with the substitution term in the expression of the elasticity of demand for one commodity with respect to the price of another. The preferred form of index number—a geometric mean with weights proportional to expenditure—is said to depend, through the weights, on income. The weights clearly also depend on prices; the differentiation which justifies the preference for this form of index is only valid if they are independent of prices.

Despite the hesitation which has been expressed in accepting a few of the authors' results, the book remains a most valuable contribution on the lognormal distribution and its applications. It also contains many useful tables and a long list of references. W. J. CORLETT

O. L. Davies (editor): *Statistical Methods in Research and Production, with special reference to the chemical industry (Third edition, revised and enlarged)*. Edinburgh and London: Oliver and Boyd, 1957. Pp. x + 396. 45s.

It is now ten years since the appearance of the first edition of this book, compiled by a team of authors from Imperial Chemical Industries Ltd. under the editorship of Dr Davies, and published under the auspices of the Company. Then statistical methods were only beginning to take their place in the industrial and scientific world and the number of textbooks was small; now statistics has 'arrived' and the number of textbooks is large, almost embarrassingly so. But the appearance of this book will be welcome to many, for the bringing up to date of a good book needs no justification. The changed climate in statistics is reflected in this edition: it is considerably longer than its predecessors, partly owing to the inclusion of new techniques and partly because some sections have been described in greater detail than was considered to be appropriate in the first edition.

The result is admirable. Although the style of presentation is similar, with appendices to the chapters where it is desired to give some material in greater detail without interrupting the main argument, this is no patched-up revision. Most of the book has been completely rewritten and many new examples have been given. I shall not comment on the introductory chapters or on the chapters dealing with frequency data and control charts, for these have been only slightly modified, but shall concentrate on the remainder.

In chapter 4, dealing with the establishment of confidence limits in simple cases, the authors make it clear that they regard these as more useful than significance tests, though the latter are explained. In other words estimation is emphasised rather than mere testing of a hypothesis. The

ffect of non-normality of data is also considered, and a distribution-free nethod for determining confidence limits for the standard deviation is iven. The next chapter is new, and deals with the choice of the number f observations and the economics of experimentation. The chapter on nalysis of variance has been considerably modified and extended to include alculation of confidence limits for the estimated variances and Tukey's nethod for calculating confidence intervals in multiple comparisons. Only he simplest experimental designs are considered, but these are quite suffiient for a book of this type. In my experience, although most scientific vorkers may admire the statistical subtleties of a complex design, they find he interpretation of such things as second-order interactions difficult. For hose who wish to cope with such things, *The Design and Analysis of Industrial Experiments*, also edited by Dr Davies, is available.

Relationships between variables occupy two chapters, with linear relationhips in one, and multiple and curvilinear relationships in the other. The reater detail has enabled a much clearer treatment to be given to such natters as the distinction between regressions and functional relationships. The chapter on sampling has been completely rewritten to place more mphasis on inspection schemes: single, double, and sequential. The treatnent of the sampling of attributes I find quite satisfactory, with the usual erminology of producer's and consumer's risks, which are stated to be the ame as errors of the first and second kinds respectively. But in the treatnent of the sequential sampling of variables, α, which is normally the error f the first kind, is stated to be the consumer's risk of accepting unsatisactory material, and β is stated to be the producer's risk of rejecting satisactory material. The correct solution is reached in the illustrative problem, ut it could have been obtained equally well if the normal terminology had een retained, and in my opinion this would have been less confusing. In his connection the modulus signs should be omitted from the equations 1·51 on page 349.

However, this is a minor blemish on an otherwise excellent book, which oncludes with a useful set of tables. The appearance of the volume is ood; such changes in type and spacing as have been made are, I think, for he better. The price is reasonable for these days. D. G. BEECH

he 1953 Pilot Population Census for the First Population Census in Sudan. Khartoum: Sudan Government Ministry for Social Affairs, Department of Statistics, 1955. Pp. 183. 20s. 6d. *First Interim Report.* Khartoum, 1956. Pp. 52. 4s. *Supplement to Interim Reports.* Khartoum, 1956. Pp. 37. 10s. 3d.

The first ever population census of Sudan was held in 1955–56 in order o provide the Sudan government with some of the basic data necessary for levelopment plans. The first results to be published are contained in the hree volumes listed.

The holding of a census for the first time, particularly in a relatively ndeveloped country like Sudan, clearly provides many difficulties, and it vas to obtain information on the type and magnitude of some of the probems likely to be encountered that a pilot census was held in 1953. This ilot census was limited to nine areas, each selected as being representative f the different types of people and modes of living to be found in Sudan. n order to draw up the sampling scheme for the main census it was necesary to have an estimate of the variability of the population and this the pilot

census provided. In addition the suitability of the questionnaire and that of the form of organisation were tested, an estimate of the likely cost was obtained, and valuable experience was gained by the enumerators.

Owing to the scattered nomadic character of a large part of the population and the unfavourable geographic conditions, the enumeration for the main census had to be spread out over a whole year. This enabled the nomadic element of the population to be enumerated at their known and accessible grazing grounds, and also meant that the whole enumeration could be carried out by relatively few enumeration teams, thus overcoming the serious lack of persons suitably qualified to act as enumerators. The census results are to be published in a series of eight *Interim Reports;* each one will cover those census areas whose results are then ready for publication and will contain fifteen basic tables giving information on sex, age, marital status, number of wives, language, tribal group or nationality, education, occupation, crude birth, death and infant mortality rates, and number of children born to women of completed fertility. In addition to the interim reports there will be a final report, which will give final figures for the whole of Sudan and contain more detailed information than that found in the interim reports.

The *Supplement to the Interim Reports* gives a brief description of the background to the census as well as details of methods and definitions used, and also gives useful comparative figures from other countries, thus forming (especially for the outsider) a short but necessary key to the interim reports, which contain no commentary.

Apart from their factual value, which must be considerable in view of the lack of data before this census, these volumes should well repay study by anyone interested in the taking of censuses, particularly in less well-developed areas. The great range of difficulties met with in the enumeration of a population that is widely scattered—the language complications, the difficulties caused by superstition, and the necessity for estimating age by physical development—are all frankly discussed in the report on the pilot census and in the supplement volume, and provide a salutary perspective to the problems which cause much discussion in censuses and sampling in more highly developed countries.

J. A. ROWNTREE

PUBLICATION RECEIVED

Torsten Sjögren and Tage Larsson: *Oligophrenia in Combination with Congenital Ichthyosis and Spastic Disorders—A Clinical and Genetic Study*. Copenhagen: Ejnar Munksgaard, 1957. Pp. 112. 25 Swedish crowns.

INDEX

Accounts, The Classification and Coding of', review, 78
Aitchison, J., 228
Anson, C. J., 223
Average, simple sequential test for, 198

Baker, A. G., 45
Barr, A., 180
Beech, D. G., 75, 160, 231
Booker, H. S., 72
Box, G. E. P., 81
Brayer, E. F., 67
Brown, J. A. C., 228
Buckland, W. R., 227
Burke, S., 159

Carden, A., 156
Cartwright, Ann, 18
Chacón, Enrique, 74
Chittleburgh, R., 56
Classification and Coding of Accounts, The', review, 78
Computer Group, London, 69
computers, electronic, some statistical applications, 102
confidence limits:
for percentage points, 214
for range of search radar, 214
continuous measurements, analysis of, 208
Corlett, W. J., 230
Cornell, F. G., 78
Coutie, G. A., 80

Dahlberg, A. O., 225
Davies, O. L., 230
Dictionary of Statistical Terms, A', review, 227
Doll, W. R. S., 72
Dornbusch, S. M., 75
Duckworth, E., 26
Dynamic Factors in Industrial Productivity', review, 76

econometrics, the scope and limitations of, 1
economic statistics in industry, 156
Edinburgh, Industrial Applications Group, 70
Educational Statistics, The Essentials of', review, 78
engineering design, the study of variability in, 133
errors in measurement, use of regression analysis to detect, 63
'Estadística, Curso de', review, 74
evolutionary operation:
a method for increasing industrial productivity, 81
'Experimental Design and its Statistical Basis', review, 73

factorial experiments, analysis and presentation of results, 45
family morbidity inquiry, information from different informants, 18
Finney, D. J., 73
Fisher, Sir R. A., 226

Gales, Kathleen, 123
'Games, The Theory of, and Linear Programming', review, 79
Gibson, Wendy M., 114, 189
Gray, P. G., 139

Ilersic, A. R., 158
'Inspection Engineer incorporating Quality Control, The', 223
interviewers, differences between experienced, 180
investigations in the factory, how statistics can help, 72

Japan, 'Reports of Statistical Application Research', 69
Johnson, R. E., 160
Jowett, G. H., 114, 189, 227

Kay, E., 26
Kemp, K. W., 208
Kendall, M. G., 227
Klein, L. R., 1

Landblom, Nellie T., 75
Lee, A., 56
linear programming in practice, 26
'Linear Programming, The Theory of Games and', review, 79
Lipton, S., 102
'Lognormal Distribution, The', review, 228
lung cancer, smoking and its effects on, 72
Melman, S., 76
Moore, P. G., 171
Morris, Doris N., 160
Morrison, S. J., 133

national income studies in the colonies, 224
'National Income Visualized', review, 225

Neale, E. P., 74
'New Zealand Official Statistics, Guide to', review, 74
Nuttall, C., 78

Oldham, P. D., 228
optical periodograph, 208

Paine, P., 156
Pearson, Karl, 154
'Population Census in Sudan, Reports', review, 231
Prest, A. R., 224
production control, a problem in, 56
productivity, a method of increasing industrial, 81
'Productivity, Dynamic Factors in Industrial', review, 76
proportion, standard error of, 67

quality control in the Midlands, 223
quality control charts, normality in, 171

Recherche Opérationelle, Revue de, 154
regression analysis:
'three-group', multiple, 189
'three-group', simple, 114
used to detect errors in measurement, 63
'Research Work, Nellie Landblom's Copybook for Beginners in', review, 75
retail prices, new index of, 72
Risk, J. M. S., 78
road accident research, statistical methods in, 156
road traffic, sampling of, 161
Roberts, H. V., 157
Rowntree, J. A., 232

sample survey, by post, 139
sampling a skew distribution, 40
sampling of road traffic, 161
Sandon, F., 79
Saunders, C. T., 74, 226
Savage, R., 156
Schmid, C. F., 75
sequential test for average, simple, 198
Smith, C. A. B., 73
'Social Statistics, A Primer of', review, 75
socio-economic class, discriminant functions, 123
standard error of a proportion, 67
Stannage, W., 63
'Statistical Formulas, Guide to Elementary', review, 160
'Statistical Methods and Scientific Inference', review, 226
'Statistical Methods in Research and Production', review, 230
'Statistics: A New Approach', review, 157
'Statistics and their Application to Commerce', review, 158
Stevens, C. F., 214

TAM service, the, 156
Tanner, J. C., 161
Tippett, L. H. C., 72, 77, 158
Tomlinson, R. C., 198
Tulse, R., 40

Vajda, S., 79
Virginia Polytechnic Institute, 69

Wallis, W. A., 157
Wilkins, L. T., 76

Printed in Great Britain at
THE KYNOCH PRESS
Birmingham